M.M. POSTNIKOV

LEÇONS DE GÉOMÉTRIE

2e semestre

ALGÈBRE LINÉAIRE ET GÉOMÉTRIE DIFFÉRENTIELLE

MIKHAÏL POSTNIKOV

LEÇONS DE GÉOMÉTRIE

IIe semestre

ALGÈBRE LINÉAIRE ET GÉOMÉTRIE DIFFÉRENTIELLE

ÉDITIONS MIR • MOSCOU

TABLE DES MATIÈRES

Préface 9

Leçon première 10

Espaces vectoriels.— Sous-espaces vectoriels.— Intersection de sous-espaces.— Enveloppes linéaires.— Somme de sous-espaces.— Dimension d'un sous-espace vectoriel.— Dimension de la somme de sous-espaces vectoriels.— Dimension d'une enveloppe linéaire.

Leçon 2 16

Théorème du rang d'une matrice.— Rang du produit de deux matrices.— Théorème de Kronecker-Capelli. — Résolution d'un système d'équations linéaires.

Leçon 3 23

Somme directe de sous-espaces vectoriels.— Décomposition d'un espace vectoriel en somme directe de sous-espaces.— Espace vectoriel quotient.— Homomorphismes d'espaces vectoriels.— Somme directe d'espaces vectoriels.

Leçon 4 30

Espace dual.— Espaces en dualité.— Bidual.— Transformation de la base duale et des coordonnées des covecteurs.— Annulateurs.— Espace des solutions d'un système d'équations linéaires homogènes.

Leçon 5 39

Annulateur de l'annulateur et annulateurs de termes directs.— Fonctionnelles bilinéaires et formes bilinéaires.— Fonctionnelles bilinéaires dans l'espace dual.— Fonctionnelles bilinéaires mixtes.— Tenseurs.

Leçon 6 48

Produit de tenseurs.— Base de l'espace des tenseurs.— Contraction.— Espace associé à une fonctionnelle multilinéaire.

Leçon 7 53

Rang d'une fonctionnelle multilinéaire.— Fonctionnelles et permutations.— Application alternée.

Leçon 8 59

Fonctionnelles multilinéaires antisymétriques.— Produit extérieur.— Algèbre de Grassmann.— Produit extérieur de covecteurs.— Décomposition des fonctionnelles antisymétriques en produit extérieur de covecteurs de base.

Leçon 9 67

Base de l'espace des fonctionnelles antisymétriques.— Formules de transformation de la base de l'espace des fonctionnelles antisymétriques.— Multivecteurs.— Rang extérieur d'une fonctionnelle antisymétrique.— Théorème du rang d'un multivecteur.— Conditions d'égalité de deux multivecteurs.

Leçon 10 75

Théorème de Cartan.— Relations de Plücker.— Coordonnées plückériennes des sous-espaces.— Plans dans l'espace affine.— Plans dans l'espace projectif et leurs coordonnées.

Leçon 11 86

Fonctionnelles bilinéaires symétriques et antisymétriques.— Matrice d'une fonctionnelle bilinéaire symétrique.— Rang d'une fonctionnelle bilinéaire.— Fonctionnelles et formes quadratiques.— Théorème de Lagrange.

Leçon 12 95

Théorème de Jacobi.— Formes quadratiques sur le corps des nombres complexes et le corps des nombres réels.— Loi d'inertie.— Fonctionnelles et formes quadratiques définies positives.

Leçon 13 102

Hypersurfaces du second degré dans l'espace projectif n-dimensionnel.— Hypersurfaces du second degré dans l'espace projectif complexe et réel-complexe.— Hypersurfaces du second degré dans l'espace affine n-dimensionnel.— Hypersurfaces du second degré dans l'espace affine complexe et réel-complexe.

Leçon 14 112

Algèbre des opérateurs linéaires.— Opérateurs et fonctionnelles bilinéaires mixtes.— Opérateurs linéaires et matrices.— Opérateurs inversibles.— Opérateur adjoint.— Alternative de Fredholm.— Sous-espaces invariants et opérateurs induits.

Leçon 15 121

Valeurs propres.— Racines caractéristiques.— Opérateurs diagonalisables.— Opérateurs à spectre simple.— Existence d'une base dans laquelle la matrice d'un opérateur est triangulaire.— Opérateurs nilpotents.

Leçon 16 129

Décomposition d'un opérateur nilpotent en somme directe d'opérateurs cycliques.— Sous-espaces principaux.— Forme normale de Jordan.— Théorème de Hamilton-Cayley.

Leçon 17 137

Complexifié d'un opérateur linéaire.— Sous-espaces propres associés aux racines caractéristiques.— Opérateurs dont les complexifiés sont diagonalisables.

Leçon 18 144

Espaces euclidiens et hermitiens.— Supplémentaires orthogonaux.— Identification de vecteurs et de covecteurs.— Annulateurs et supplémentaires orthogonaux.— Fonctionnelles bilinéaires et opérateurs linéaires.— Une règle imposée à l'identification de tenseurs de différents types.— Tenseur métrique.— Elévation et abaissement d'indices.

Leçon 19 154

Opérateurs adjoints.— Opérateurs auto-adjoints.— Opérateurs antisymétriques et antihermitiens.— Une analogie entre les opérateurs hermitiens et les nombres réels.— Propriétés spectrales des opérateurs auto-adjoints.— Propriété des opérateurs auto-adjoints d'être orthogonalement diagonalisables.

Leçon 20 161

Réduction canonique des formes quadratiques par une transformation orthogonale de variables.— Hypersurfaces du second degré dans l'espace ponctuel euclidien.— Propriété de maximum-minimum des valeurs propres des opérateurs auto-adjoints.— Opérateurs orthogonalement diagonalisables.

Leçon 21 169

Opérateurs positifs.— Opérateurs isométriques.— Matrices unitaires.— Décomposition polaire des opérateurs inversibles.— Interprétation géométrique de la décomposition polaire.— Translations et transformations affines centrées.— Diagonalisation d'un opérateur unitaire.— Décomposition de la rotation de l'espace euclidien de dimension n en produit de rotations de plans bidimensionnels.

Leçon 22 179

Fonctions différentiables.— Hypersurfaces différentiables.— Gradient.— Dérivées par rapport à un vecteur.— Champs vectoriels.— Points singuliers des champs vectoriels.— Champs vectoriels en tant que modules.— Champs de potentiel et champs irrotationnels.— Rotationnel.— Divergence des champs vectoriels.— Analyse vectorielle.— Hamiltonien.— Formules de produits.— Opérateurs composés.

Leçon 23 197

Courbes continues, différentiables et régulières.—Courbes équivalentes.— Courbes régulières dans le plan et graphes des fonctions.— Hyperplan tangent à une hypersurface.— Longueur d'une courbe.— Courbes dans le plan.— Courbes dans l'espace de dimension 3.

Leçon 24 . . . . . 21[illegible]

Projections d'une courbe sur les plans de coordonnées du repère mobile.— Formules de Frenet pour une courbe dans l'espace n-dimen-

sionnel.— Une courbe définie par ses courbures.— Surfaces régulières.— Exemples de surfaces.

Leçon 25 226

Vecteurs tangents à une surface.— Plan tangent.— Première forme quadratique d'une surface.— Longueurs et angles sur une surface.— Difféomorphismes de surfaces.— Isométrie et géométrie intrinsèque d'une surface.— Exemples.— Surfaces développables.

Leçon 26 239

Plan tangent et vecteur de la normale.— Courbure d'une section normale.— Deuxième forme quadratique d'une surface.— Indicatrice de Dupin.— Courbure principales.— Deuxième forme quadratique d'un graphe.— Surfaces réglées à courbure nulle.— Surfaces de révolution.

Leçon 27 255

Formules de dérivation de Weingarten. — Coefficients de connexion. — Théorème de Gauss. — Conditions nécessaires et suffisantes de surfaces isométriques.

Index 260

PRÉFACE

Ce livre continue la *Géométrie analytique* *) et reproduit presque textuellement les conférences faites par l'auteur (dans le cadre du cours d'Algèbre linéaire et de Géométrie analytique) à l'intention des étudiants de première année de la Faculté mécanico-mathématique de l'Université de Moscou. Le choix des matières et leur succession sont naturellement dictés par les mêmes raisons qu'au cours du premier semestre (voir I, préface à l'édition russe). Le nombre de leçons est celui qu'on lit pratiquement 27 encore que le programme en prévoie 32.

Le cours d'Algèbre linéaire et de Géométrie analytique fait partie du cours biennal de Géométrie, si bien que le choix des thèmes et leur importance relative sont pour beaucoup déterminés par les besoins de la deuxième année scolaire consacrée à la Géométrie différentielle des variétés. Ainsi, l'auteur a incorporé dans ce livre (en marge du programme) certaines matières propédeutiques du troisième semestre (la Géométrie différentielle élémentaire des courbes et des surfaces dans l'espace de dimension 3), dont l'approche se trouve donc sensiblement facilitée (et ce, non seulement au conférencier, mais aussi, chose essentielle, aux étudiants). L'expérience a d'ailleurs montré que ces matières présentent de l'intérêt pour les étudiants qui les assimilent bien dès le deuxième semestre.

27 octobre 1977 *Mikhaïl Postnikov*

*) M. Postnikov, *Leçons de géométrie*: Semestre premier. Géométrie analytique, Ed. de Moscou, 1981. Dans la suite, on désigne ce livre par I. (Le chiffre arabe qui suit I désigne le numéro de la leçon. (*N.d.T.*))

LEÇON PREMIÈRE

Espaces vectoriels. — Sous-espaces vectoriels. — Intersection de sous-espaces. — Enveloppes linéaires. — Somme de sous-espaces. — Dimension d'un sous-espace vectoriel. — Dimension de la somme de sous-espaces vectoriels. — Dimension d'une enveloppe linéaire.

Nous allons généraliser les résultats du semestre précédent à n quelconque. Les matières seront pour l'essentiel présentées dans l'esprit du Livre I *).

On rappelle (voir I.1, déf. 1) qu'un *espace vectoriel* sur un corps $\mathbb{K}$ est par définition un ensemble $\mathcal{V}$ d'éléments appelés *vecteurs* muni d'une opération d'addition $\mathbf{x}, \mathbf{y} \mapsto \mathbf{x} + \mathbf{y}$ et d'une opération de multiplication $\mathbf{x} \mapsto k\mathbf{x}$ pour tout $k \in \mathbb{K}$ telles que l'ensemble $\mathcal{V}$ soit un groupe abélien pour l'addition et que l'opération de multiplication remplisse quatre axiomes naturels.

Dans cet espace, la notion de combinaison linéaire de vecteurs a un sens, ainsi que les notions de familles et de systèmes libres et liés. Un espace $\mathcal{V}$ est dit *de dimension finie* s'il existe une *base* finie de $\mathcal{V}$, i.e. une famille de vecteurs tels que tout vecteur de $\mathcal{V}$ soit représentable d'une façon unique comme leur combinaison linéaire. Toutes les bases ont le même nombre d'éléments. Ce nombre s'appelle la dimension de l'espace vectoriel $\mathcal{V}$ et se note $\dim \mathcal{V}$.

* * *

Soit $\mathcal{V}$ un espace vectoriel quelconque de dimension finie.

Définition 1. Une partie $\mathcal{P}$ de $\mathcal{V}$ est un *sous-espace vectoriel* de $\mathcal{V}$ si toute combinaison linéaire $k_1\mathbf{x}_1 + \ldots + k_m\mathbf{x}_m$ de $\mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathcal{P}$ quelconques appartient à $\mathcal{P}$.

Une condition nécessaire et suffisante pour que $\mathcal{P}$ soit un sous-espace vectoriel est évidemment que $\mathbf{x} + \mathbf{y} \in \mathcal{P}$ et $k\mathbf{x} \in \mathcal{P}$ pour $\mathbf{x}, \mathbf{y} \in \mathcal{P}$ quelconques et tout $k \in \mathbb{K}$.

Autrement dit, la propriété, pour $\mathcal{P}$, d'être un sous-espace vectoriel signifie que les correspondances $\mathbf{x}, \mathbf{y} \mapsto \mathbf{x} + \mathbf{y}$ et $\mathbf{x} \mapsto k\mathbf{x}$, avec $\mathbf{x}, \mathbf{y} \in \mathcal{P}$ et $k \in \mathbb{K}$, définissent sur $\mathcal{P}$ certaines opérations. Il est clair que le *sous-espace $\mathcal{P}$ est un espace vectoriel* par rapport à ces opérations. □

---

*) Comme dans I, les notations et les abréviations sont celles de l'auteur. (*N.d.T.*)

Exemples de sous-espaces vectoriels.

1. Pour tout espace vectoriel $\mathcal{V}$, la partie de $\mathcal{V}$ réduite au vecteur nul et l'espace vectoriel $\mathcal{V}$ tout entier sont des sous-espaces vectoriels. Le singleton $\{0\}$ (noté également 0) s'appelle le sous-espace *nul*, et le sous-espace $\mathcal{V}$ est dit *trivial*.

2. Pour l'espace vectoriel $\mathbb{K}^n$ et $m \leqslant n$ quelconque, l'ensemble des vecteurs $(x^1, \ldots, x^m, 0, \ldots, 0)$ de $\mathbb{K}^n$ ayant leurs $n - m$ dernières coordonnées nulles est un sous-espace vectoriel. Ce sous-espace est canoniquement isomorphe à l'espace $\mathbb{K}^m$.

3. Dans l'espace vectoriel des polynômes (ou, de façon plus générale, des fonctions quelconques vérifiant certaines conditions), l'ensemble des polynômes (des fonctions) nuls en un ou en plusieurs points fixes constitue un sous-espace vectoriel.

4. Tous les polynômes dont les puissances fixes données ont leurs coefficients nuls forment un sous-espace vectoriel, ainsi que tous les polynômes pairs ou tous les polynômes impairs.

* * *

Proposition 1. *L'intersection*

$$\mathcal{P} = \bigcap_{\alpha} \mathcal{P}_\alpha$$

d'une famille quelconque de sous-espaces $\mathcal{P}_\alpha \subset \mathcal{V}$ est un sous-espace de $\mathcal{V}$.

Démonstration. Si $\mathbf{x}, \mathbf{y} \in \mathcal{P}$, alors $\mathbf{x}, \mathbf{y} \in \mathcal{P}_\alpha$ pour tout α, si bien que $\mathbf{x} + \mathbf{y} \in \mathcal{P}_\alpha$, $k\mathbf{x} \in \mathcal{P}_\alpha$, donc $\mathbf{x} + \mathbf{y} \in \mathcal{P}$, $k\mathbf{x} \in \mathcal{P}$ (α étant arbitraire). □

On note que l'intersection de sous-espaces vectoriels ne peut être vide du moment que tout sous-espace contient le vecteur nul 0.

Si $\mathcal{P} \cap \mathcal{Q} = 0$, les sous-espaces $\mathcal{P}$ et $\mathcal{Q}$ sont dits *disjoints*.

Bien que simple, la proposition 1 entraîne des résultats importants.

* * *

Soit S une partie quelconque de l'espace vectoriel $\mathcal{V}$.

Définition 2. On dit qu'un sous-espace $\mathcal{P} \subset \mathcal{V}$ constitue l'*enveloppe linéaire* de l'ensemble S (ou qu'il est *engendré* par S) si $S \subset \mathcal{P}$ et $\mathcal{P}$ est le plus petit sous-espace vectoriel de $\mathcal{V}$ ayant cette propriété, i.e. si chaque sous-espace $\mathcal{Q}$ tel que $S \subset \mathcal{Q}$ contient $\mathcal{P}$. On note $[S]$ l'enveloppe linéaire de l'ensemble S.

Proposition 2. *Tout ensemble $S \subset \mathcal{V}$ admet une enveloppe linéaire $[S]$ qui est l'intersection de tous les sous-espaces contenant S.*

Démonstration. Comme cette intersection (qui est un sous-espace

vectoriel par la proposition 1) porte sur tout sous-espace $\mathcal{Q} \supset S$, elle est contenue dans $\mathcal{Q}$. D'autre part, elle contient évidemment S. □

Une question se pose : peut-on en général parler de l'intersection de sous-espaces contenant S ? et ces sous-espaces existent-ils ? La réponse formelle est la suivante. Selon la théorie générale des ensembles, l'intersection d'une famille de parties d'un ensemble $\mathcal{V}$ quelconque est définie même si la famille est vide, et, chose paradoxale, elle est dans ce cas $\mathcal{V}$ tout entier. Or, la famille considérée n'est jamais vide. En effet, un des sous-espaces contenant S est nécessairement l'espace $\mathcal{V}$ tout entier.

On se représente mieux l'enveloppe linéaire $[S]$ grâce à la

Proposition 3. *L'enveloppe linéaire* $[S]$ *de l'ensemble* S *est constituée par toutes les combinaisons linéaires*

$$(1) \qquad k_1\mathbf{x}_1 + \ldots + k_m\mathbf{x}_m, \quad \mathbf{x}_1, \ldots, \mathbf{x}_m \in S, \quad k_1, \ldots, k_m \in \mathbb{K},$$

des vecteurs de S.

Démonstration. Si $\mathcal{P}$ est un sous-espace contenant S, il renferme évidemment tous les vecteurs de la forme (1). D'autre part, l'ensemble des vecteurs (1) est évidemment un sous-espace contenant S. □

Cette proposition entraîne qu'*un ensemble de vecteurs de* $\mathcal{V}$ *est générateur si et seulement si le sous-espace vectoriel qu'il engendre est égal à* $\mathcal{V}$ *tout entier.* □

On rappelle (I.12) que deux ensembles de vecteurs sont dits *linéairement équivalents* si chaque vecteur d'un ensemble s'exprime linéairement au moyen des vecteurs de l'autre. Cela équivaut à dire qu'un vecteur est combinaison linéaire des vecteurs de l'un des ensembles si et seulement s'il est combinaison linéaire des vecteurs de l'autre, i.e. que, selon la proposition 3, les *enveloppes linéaires de deux ensembles coïncident* (les deux ensembles engendrent un même sous-espace).

* * *

Contrairement à l'intersection, la réunion de sous-espaces vectoriels n'est en général pas un sous-espace vectoriel. On obtient un sous-espace par passage à l'enveloppe linéaire de la réunion.

Définition 3. Etant donnée une famille arbitraire de sous-espaces $\mathcal{P}_\alpha \subset \mathcal{V}$, le sous-espace vectoriel engendré par leur réunion s'appelle la *somme* de la famille et se note

$$\sum_\alpha \mathcal{P}_\alpha = [\bigcup_\alpha \mathcal{P}_\alpha].$$

Pour deux sous-espaces $\mathcal{P}$ et $\mathcal{Q}$ on a

$$\mathcal{P} + \mathcal{Q} = [\mathcal{P} \cup \mathcal{Q}].$$

Il est clair que toute combinaison linéaire des vecteurs de $\mathcal{P} \cup \mathcal{Q}$ s'écrit $\mathbf{x} + \mathbf{y}$, où $\mathbf{x} \in \mathcal{P}$, $\mathbf{y} \in \mathcal{Q}$, ce qui démontre la

Proposition 4. *La somme* $\mathcal{P} + \mathcal{Q}$ *des sous-espaces* $\mathcal{P}$ *et* $\mathcal{Q}$ *est formée de tous les vecteurs de la forme* $\mathbf{x} + \mathbf{y}$, *avec* $\mathbf{x} \in \mathcal{P}$, $\mathbf{y} \in \mathcal{Q}$. □

On a une proposition analogue pour toute famille de sous-espaces vectoriels.

* * *

Nous n'avons pas une seule fois utilisé l'hypothèse de $\mathcal{V}$ de dimension finie. Cette hypothèse est essentielle pour les résultats ci-dessous.

Soit $n = \dim \mathcal{V}$.

Proposition 5. *Etant donné un sous-espace* $\mathcal{P} \subset \mathcal{V}$ *quelconque, on a l'inégalité*

$$\dim \mathcal{P} \leqslant n.$$

Des étudiants (voire des auteurs) le démontrent de façon suivante: n'importe quels $n + 1$ vecteurs de $\mathcal{P}$ étant éléments de l'espace $\mathcal{V}$ n-dimensionnel sont linéairement dépendants; le sous-espace $\mathcal{P}$ contient donc n vecteurs linéairement indépendants au plus, ce qui donne $\dim \mathcal{P} \leqslant n$.

Le raisonnement est déficient parce que $\mathcal{P}$ est supposé de dimension finie, ce qui ne découle nullement de façon directe de la propriété analogue de $\mathcal{V}$. On n'établit donc en fait qu'une base de $\mathcal{P}$, quand elle existe, contient n éléments au plus. On aboutit au résultat 5 par des raisonnements plus subtiles.

Démonstration de la proposition 5. Si $\mathcal{P} = 0$, le résultat est évident. Si $\mathcal{P} \neq 0$, il existe un vecteur $\mathbf{e}_1 \in \mathcal{P}$ non nul. Lorsque $\mathcal{P} = [\mathbf{e}_1]$, le vecteur $\mathbf{e}_1$ constitue évidemment une base de $\mathcal{P}$, donc $\dim \mathcal{P} = 1$. Soit $\mathcal{P} \neq [\mathbf{e}_1]$. On trouve dans $\mathcal{P}$ un vecteur $\mathbf{e}_2$ qui ne s'exprime pas linéairement par $\mathbf{e}_1$, i.e. tel que $\mathbf{e}_1$, $\mathbf{e}_2$ soient linéairement indépendants. Si $\mathcal{P} = [\mathbf{e}_1, \mathbf{e}_2]$, alors $\mathbf{e}_1$, $\mathbf{e}_2$ sont une base de $\mathcal{P}$, et $\dim \mathcal{P} = 2$. Si $\mathcal{P} \neq [\mathbf{e}_1, \mathbf{e}_2]$, alors il existe dans $\mathcal{P}$ un ve teur $\mathbf{e}_3$ qui n'est pas combinaison linéaire de $\mathbf{e}_1$, $\mathbf{e}_2$, et ainsi de suite. Comme $\dim \mathcal{V} = n$, le processus s'arrête avec l'apparition de $\mathbf{e}_n$, si bien que $\mathcal{P}$ est de dimension finie et $\dim \mathcal{P} \leqslant n$. □

Si $\dim \mathcal{P} = n$, toute base de $\mathcal{P}$ étant une famille libre de n vecteurs est une base de $\mathcal{V}$. Aussi $\mathcal{P} = \mathcal{V}$. Si $\dim \mathcal{P} < n$, toute base de $\mathcal{P}$ ayant moins de n éléments n'est donc pas une famille génératrice de $\mathcal{V}$ et n'engendre pas $\mathcal{V}$ tout entier. Aussi $\mathcal{P} \neq \mathcal{V}$. Ainsi, *un sous-espace* $\mathcal{P} \subset \mathcal{V}$ *est égal à* $\mathcal{V}$ *si et seulement si* $\dim \mathcal{P} = \dim \mathcal{V}$. □

* * *

Théorème 1 (dimension de la somme de deux sous-espaces vectoriels). *Etant donnés deux sous-espaces vectoriels* $\mathcal{P}$ *et* $\mathcal{Q}$ *quelconques, on a la formule*

$$\dim (\mathcal{P} + \mathcal{Q}) = \dim \mathcal{P} + \dim \mathcal{Q} - \dim (\mathcal{P} \cap \mathcal{Q}).$$

Démonstration. Soit

$$\dim \mathcal{P} = p, \quad \dim \mathcal{Q} = q, \quad \dim (\mathcal{P} \cap \mathcal{Q}) = r$$

et soit, dans $\mathcal{P} \cap \mathcal{Q}$, une base $\mathbf{e}_1, \ldots, \mathbf{e}_r$ arbitraire. En la complétant, on a finalement une base

$$(2) \qquad \mathbf{e}_1, \ldots, \mathbf{e}_r, \quad \mathbf{f}_1, \ldots, \mathbf{f}_{p-r}$$

du sous-espace $\mathcal{P} \supset \mathcal{P} \cap \mathcal{Q}$. S'agissant de $\mathcal{Q}$, on obtient de même une base

$$(3) \qquad \mathbf{e}_1, \ldots, \mathbf{e}_r, \quad \mathbf{g}_1, \ldots, \mathbf{g}_{q-r}.$$

On a manifestement le résultat cherché si l'on démontre que $p + q - r$ vecteurs

$$(4) \qquad \mathbf{e}_1, \ldots, \mathbf{e}_r, \quad \mathbf{f}_1, \ldots, \mathbf{f}_{p-r}, \quad \mathbf{g}_1, \ldots, \mathbf{g}_{q-r}$$

constituent une base du sous-espace $\mathcal{P} + \mathcal{Q}$.

Indépendance linéaire. Soit

$$k_1\mathbf{e}_1 + \ldots + k_r\mathbf{e}_r + l_1\mathbf{f}_1 + \ldots + l_{p-r}\mathbf{f}_{p-r} + m_1\mathbf{g}_1 + \ldots + m_{q-r}\mathbf{g}_{q-r} = \mathbf{0}.$$

On pose

$$\mathbf{e} = k_1\mathbf{e}_1 + \ldots + k_r\mathbf{e}_r,$$
$$\mathbf{f} = l_1\mathbf{f}_1 + \ldots + l_{p-r}\mathbf{f}_{p-r},$$
$$\mathbf{g} = m_1\mathbf{g}_1 + \ldots + m_{q-r}\mathbf{g}_{q-r},$$

et on obtient $\mathbf{e} \in \mathcal{P} \cap \mathcal{Q}$, $\mathbf{f} \in \mathcal{P}$ et $\mathbf{g} \in \mathcal{Q}$ tels que $\mathbf{e} + \mathbf{f} + \mathbf{g} = \mathbf{0}$, auquel cas $\mathbf{e} + \mathbf{f} \in \mathcal{P}$, donc $\mathbf{g} = -(\mathbf{e} + \mathbf{f}) \in \mathcal{P}$. Ainsi, $\mathbf{g} \in \mathcal{P} \cap \mathcal{Q}$. ce qui entraîne que $\mathbf{g}$ s'exprime linéairement par $\mathbf{e}_1, \ldots, \mathbf{e}_r$ Mais le vecteur $\mathbf{g}$ est par hypothèse une combinaison linéaire de $\mathbf{g}_1, \ldots, \mathbf{g}_{p-r}$. Un vecteur est représentable de façon unique par la base (3), ce qui démontre que les deux expressions de $\mathbf{g}$ ont leurs coefficients nuls. Ainsi, $m_1 = 0, \ldots, m_{q-r} = 0$, donc $\mathbf{g} = \mathbf{0}$.

Dans ce cas, $\mathbf{e} + \mathbf{f} = \mathbf{0}$ et $k_1 = 0, \ldots, k_r = 0, l_1 = 0, \ldots, l_{p-r} = 0$ (vu que (2) est une base), i.e. les vecteurs (4) sont linéairement indépendants, c.q.f.d.

Famille génératrice. On sait que tout vecteur de $\mathcal{P} + \mathcal{Q}$ s'écrit $\mathbf{x} + \mathbf{y}$, où $\mathbf{x} \in \mathcal{P}$, $\mathbf{y} \in \mathcal{Q}$. La décomposition de $\mathbf{x}$ dans la base (2) plus la décomposition de $\mathbf{y}$ dans la base (3) fournit $\mathbf{x} + \mathbf{y}$ comme combinaison linéaire des vecteurs (4), i.e. la famille (4) de vecteurs de $\mathcal{P} + \mathcal{Q}$ est génératrice.

La famille (4) étant libre et génératrice constitue une base. □

Conséquence 1. *Si* $\mathcal{P} + \mathcal{Q} = \mathcal{V}$, *alors* $\dim (\mathcal{P} \cap \mathcal{Q}) = p + q - n$.

Conséquence 2. *Si* $p + q > n$, *alors* $\mathcal{P} \cap \mathcal{Q} \neq 0$.

* * *

Il y a autant de procédés de calcul de la dimension d'un sous-espace vectoriel que de façons dont le sous-espace est donné, si bien que nous reviendrons constamment sur cette question. Le seul procédé constructif pour se donner un sous-espace vectoriel que nous connaissons en fait pour l'instant est de le définir comme enveloppe linéaire d'un ensemble fini de vecteurs. Le problème général se concrétise donc dans la recherche de la dimension de l'enveloppe linéaire $[S]$ d'un ensemble S (fini) quelconque de vecteurs.

Soit S un tel ensemble. On estime sans restreindre la généralité que ses éléments sont non tous nuls et qu'il contient donc des parties libres. Les vecteurs de S étant en nombre fini, certaines parties sont *maximales*, i.e. elles deviennent des ensembles liés par adjonction de tout autre vecteur de S. La chose n'est possible que si le vecteur ajouté est combinaison linéaire des éléments de la partie, ce qui entraîne que *toute partie libre maximale S_0 de l'ensemble S et l'ensemble S sont linéairement équivalents*, i.e. S_0 engendre le même sous-espace $[S]$ (voir plus haut). Cela veut dire que S_0 est générateur dans $[S]$. Comme S_0 est de plus libre, un numérotage arbitraire en fait une base de $[S]$. Ainsi, *toute sous-famille libre maximale de l'ensemble S est une base de l'enveloppe linéaire $[S]$ de S.*

Puisque toutes les bases d'un espace ont le même nombre d'éléments, il en résulte en particulier que *toutes les parties libres maximales de l'ensemble S ont le même nombre de vecteurs.*

Définition 4. Le nombre de vecteurs d'une partie libre maximale de l'ensemble S s'appelle le *rang* de S.

La définition est intrinsèque (on vient de l'établir).

On a en outre la

Proposition 6. *La dimension* dim $[S]$ *de l'enveloppe linéaire de l'ensemble de vecteurs S est égale au rang de S.* □

On dirait une tautologie peu intéressante. Rien n'est plus faux. Il s'agit en fait d'un résultat très important qui identifie dim $[S]$ cherché et le rang qu'on trouve, ne serait-ce que théoriquement, après n pas, n étant évalué à l'avance, i.e. qui est un nombre *effectivement calculable*. En effet, un procédé de recherche du rang consiste par exemple à explorer successivement toutes les parties de S (qui sont en nombre fini !) et à vérifier si telle partie est oui ou non libre (ce qui s'effectue également en un nombre fini de pas). Ainsi, l'intérêt de la proposition 6 est qu'elle indique une méthode finie de calcul de la dimension des sous-espaces vectoriels *si* ces derniers sont définis comme étant les enveloppes linéaires des ensembles finis de vecteurs (la condition d'ensembles finis est obligatoire pour que la méthode soit efficace !).

Le volume de calcul est sensiblement diminué par une conduite raisonnable des opérations. La technique correspondante sera étudiée dans la leçon 2.

LEÇON 2

Théorème du rang d'une matrice. — Rang du produit de deux matrices. — Théorème de Kronecker-Capelli. — Résolution d'un système d'équations linéaires.

A la fin de la leçon précédente nous avons parlé de la possibilité de calculer rationnellement le rang d'un ensemble de vecteurs. La méthode dépend naturellement de la façon dont on se donne les vecteurs. Dans le seul procédé que nous nous proposons d'examiner (qui est d'ailleurs le plus important), les vecteurs sont définis par leurs coordonnées dans une base. Cela équivaut à les supposer dans l'espace K^n des vecteurs lignes.

Etant donnés m vecteurs

$$(1)\qquad \begin{aligned} \mathbf{a}_1 &= (a_{11}, \ldots, a_{1n}), \\ \mathbf{a}_m &= (a_{m1}, \ldots, a_{mn}) \end{aligned}$$

de l'espace K^n, on forme avec leurs composantes la matrice rectangulaire

$$(2)\qquad A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ a_{m1} & \ldots & a_{mn} \end{pmatrix},$$

et on énonce le problème posé en d'autres termes, à savoir « soit la matrice rectangulaire (2). On demande le rang de l'ensemble des lignes de (2) ». C'est sous cette forme que nous allons le résoudre.

Soit $1 \leqslant p \leqslant \min(m, n)$. Prenons p lignes et p colonnes quelconques de la matrice A. Les éléments au croisement de ces lignes et ces colonnes forment une « sous-matrice » carrée d'ordre p. Les déterminants de ces sous-matrices s'appellent les *mineurs d'ordre p* de la matrice A.

Définition 1. On appelle *rang* de la matrice A le plus grand ordre des mineurs non nuls, i.e. un nombre p tel que tout mineur d'ordre $p + 1$ de A soit nul et que A possède un mineur d'ordre p qui ne l'est pas.

On observe que si tous les mineurs d'ordre $p + 1$ sont égaux à 0, il en est de même des mineurs d'ordre $p + 2$ (car ils sont, par la formule du développement des déterminants, des combinaisons linéaires des mineurs d'ordre $p + 1$) et des mineurs d'ordre supérieur.

Il est clair que le rang de la matrice (2) vérifie les inégalités

$$0 \leqslant p \leqslant \min (m, n),$$

p étant nul si et seulement si la matrice n'a pas d'autres éléments à part 0.

Par exploration successive des mineurs d'ordre toujours plus élevé on obtient le rang d'une matrice quelconque en un nombre fini de pas. Aussi le problème proposé est résolu par le

Théorème 1 (rang d'une matrice). *Le rang de toute matrice est égal au rang de l'ensemble de ses lignes.*

Démonstration. On commence par noter que tout échange des lignes ou des colonnes de A établit une bijection entre l'ensemble de tous ses mineurs de chaque ordre et l'ensemble des mineurs de même ordre de la matrice transformée, les mineurs non nuls étant changés en des mineurs non nuls. Par conséquent, tout échange conserve le rang p de la matrice A.

Et le rang des lignes? Il est évidemment conservé par permutation des lignes. Quant à l'échange des colonnes, elle se ramène à changer les notations des composantes de tous les vecteurs (1), si bien que les relations linéaires entre ces vecteurs (ou entre quelques-uns d'eux) restent manifestement entières. Toute permutation des colonnes de A conserve donc le rang de l'ensemble de ses lignes.

Par échange des lignes et des colonnes on obtient une matrice A telle que son mineur d'ordre p non nul soit dans son angle supérieur gauche. Il en résulte que l'égalité $p = r$ est démontrée en supposant (sans restreindre la généralité)

$$\Delta = \begin{vmatrix} a_{11} & \dots & a_{1p} \\ \dots & \dots & \dots \\ a_{p1} & \dots & a_{pp} \end{vmatrix} \neq 0.$$

Si p premières lignes de la matrice A étaient linéairement dépendantes, il en serait de même des lignes du déterminant Δ, d'où la nullité de Δ. Ainsi, *les lignes* $\mathbf{a}_1, \dots, \mathbf{a}_p$ *de la matrice A sont linéairement indépendantes, donc* $p \leqslant r$.

On établit donc $p = r$ sous la condition suffisante que *chaque ligne* $\mathbf{a}_i$, $i > p$, *s'exprime linéairement au moyen des lignes* $\mathbf{a}_1, \dots, \mathbf{a}_p$.

On considère le déterminant d'ordre $p + 1$:

$$\begin{vmatrix} a_{11} & \dots & a_{1p} & a_{1j} \\ a_{21} & \dots & a_{2p} & a_{2j} \\ \dots & \dots & \dots & \dots \\ a_{p1} & \dots & a_{pp} & a_{pj} \\ a_{i1} & \dots & a_{ip} & a_{ij} \end{vmatrix}, \tag{3}$$

où $1 \leqslant j \leqslant n$. Si $1 \leqslant j \leqslant p$, le déterminant (3) a deux colonnes identiques et est donc nul. Si $p + 1 \leqslant j \leqslant n$, alors (3) est le mineur d'ordre $p + 1$ de A (qui s'obtient en choisissant p premières lignes et colonnes, la colonne j et la ligne i) et est nul lui aussi. En le développant suivant la dernière colonne, on a pour tout $j = 1, \ldots, n$

$$A_1 a_{1j} + A_2 a_{2j} + \ldots + A_p a_{pj} + \Delta a_{ij} = 0, \tag{4}$$

où $A_1, A_2, \ldots, A_p, \Delta$ sont les cofacteurs des éléments de la colonne. Les cofacteurs dépendent seulement des éléments des p premières colonnes de (3), et ils sont en particulier égaux pour tous les j. En notations vectorielles, n égalités (4) sont donc équivalentes à une égalité de la forme

$$A_1 \mathbf{a}_1 + A_2 \mathbf{a}_2 + \ldots + A_p \mathbf{a}_p + \Delta \mathbf{a}_i = 0,$$

ce qui démontre ($\Delta \neq 0$ par hypothèse) que le vecteur $\mathbf{a}_i$, $p + 1 \leqslant i \leqslant n$, est combinaison linéaire des vecteurs $\mathbf{a}_1, \ldots, \mathbf{a}_p$. Par conséquent, $r = p$. □

La démonstration du théorème 1 entraîne en particulier que *si la matrice A a un mineur d'ordre p non nul tel que tous les mineurs d'ordre $p + 1$ bordants soient nuls, alors la matrice est de rang p.* □

Cela simplifie sensiblement la recherche du rang.

Dans le cas particulier de A carrée d'ordre égal à son rang, on a la

Conséquence. *Un déterminant est non nul si et seulement si ses lignes sont linéairement indépendantes.* □

Il est clair que A et sa transposée ont même rang p. D'autre part, le rang des colonnes de A est égal au rang des lignes de sa transposée, i.e. *le rang de l'ensemble des lignes d'une matrice quelconque est égal au rang de l'ensemble de ses colonnes.* □

N'est-ce pas un résultat intéressant qui établit une relation entre les rangs des familles de vecteurs de deux espaces vectoriels sans parler du fait que ces espaces sont en général de dimensions différentes ? !

* * *

Que devient le rang d'une matrice multipliée par une autre ?

Soient A une matrice à n colonnes et m lignes (cas ci-dessus) et B une matrice à n lignes et s colonnes. On définit alors la matrice AB à m lignes et s colonnes. Si $r(A)$ est le rang de A et $r(B)$ le rang de B, que peut-on dire du rang $r(AB)$ de leur produit ?

Il se trouve qu'on n'en sait rien dans le cas général sinon que $r(AB)$ est au plus égal au plus petit des $r(A)$ et $r(B)$.

Proposition 1. *On a les inégalités*

$$r(AB) \leqslant r(A), \quad r(AB) \leqslant r(B).$$

Démonstration. Soient

$$A=\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad B=\begin{pmatrix} b_{11} & \dots & b_{1s} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{ns} \end{pmatrix}, \quad AB=\begin{pmatrix} c_{11} & \dots & c_{1s} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{ms} \end{pmatrix}.$$

Par définition du produit de matrices

$$c_{ik}=\sum_{j=1}^{n} a_{ij}b_{jk}, \quad i=1, \dots, m, \quad k=1, \dots, s.$$

On introduit les vecteurs lignes des matrices B et C:

$$\begin{array}{ll} \mathbf{b}_1 = (b_{11}, \dots, b_{1s}), & \mathbf{c}_1 = (c_{11}, \dots, c_{1s}), \\ \dots\dots\dots\dots & \dots\dots\dots\dots \\ \mathbf{b}_n = (b_{n1}, \dots, b_{ns}), & \mathbf{c}_m = (c_{m1}, \dots, c_{ms}). \end{array}$$

Les formules de c_{ik} se récrivent:

$$\mathbf{c}_i=\sum_{j=1}^{n} a_{ij}\mathbf{b}_j, \quad i=1, \dots, m,$$

i.e. les vecteurs $\mathbf{c}_1, \dots, \mathbf{c}_m$ s'expriment linéairement en fonction de $\mathbf{b}_1, \dots, \mathbf{b}_n$. Par conséquent,

$$[\mathbf{c}_1, \dots, \mathbf{c}_m] \subset [\mathbf{b}_1, \dots, \mathbf{b}_n],$$

donc

$$\dim [\mathbf{c}_1, \dots, \mathbf{c}_m] \leqslant \dim [\mathbf{b}_1, \dots, \mathbf{b}_n],$$

i.e. $r(AB) \leqslant r(B)$ par le théorème du rang d'une matrice.

L'inégalité $r(AB) \leqslant r(A)$ se démontre de façon analogue (à condition de remplacer les lignes par les colonnes). On peut également utiliser l'inégalité démontrée, la propriété, pour le rang, d'être conservé par transposition et l'égalité $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$. En effet,

$$r(AB) = r((AB)^{\mathsf{T}}) = r(B^{\mathsf{T}}A^{\mathsf{T}}) \leqslant r(A^{\mathsf{T}}) = r(A). \quad \square$$

Si A ou B est carrée non dégénérée, on démontre un résultat plus précis, à savoir on a la

Proposition 2. *Si B est une matrice carrée ($n = s$) non dégénérée* (det $B \neq 0$), *alors on a pour toute matrice A:*

$$r(AB) = r(A).$$

Si A est une matrice carrée ($n = m$) non dégénérée (det $A \neq 0$), *alors on a pour toute matrice B:*

$$r(AB) = r(B).$$

Bref, la multiplication d'une matrice par une matrice non dégénérée conserve le rang.

Démonstration. Pour B non dégénérée, il existe la matrice inverse B^{-1}, et on a $A = (AB)^{\bullet}B^{-1}$. Selon la proposition 1,

$$r(A) = r((AB)B^{-1}) \leqslant r(AB).$$

Donc, $r(A) = r(AB)$. On démontre de façon analogue $r(B) = r(AB)$ pour A non dégénérée. □

* * *

Le théorème du rang d'une matrice permet non seulement de calculer effectivement les rangs et de trouver les parties libres maximales, mais encore de vérifier par exemple si le vecteur donné $\mathbf{b}$ s'exprime par $\mathbf{a}_1, \ldots, \mathbf{a}_m$ donnés (sans chercher sous forme explicite les coefficients de la relation linéaire).

En effet, *le vecteur* $\mathbf{b}$ *s'exprime de façon linéaire au moyen des vecteurs* $\mathbf{a}_1, \ldots, \mathbf{a}_m$ *si et seulement si* chaque partie libre maximale de l'ensemble $\mathbf{a}_1, \ldots, \mathbf{a}_m$ en est une de l'ensemble augmenté $\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}$ et, partant, *si* $\mathbf{a}_1, \ldots, \mathbf{a}_m$ *et* $\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}$ *ont même rang.* □

Il y a intérêt à énoncer ce résultat en termes d'équations linéaires. Si

$$\begin{aligned} \mathbf{a}_1 &= (a_{11}, \ldots, a_{1n}), \\ &\ldots\ldots\ldots \\ \mathbf{a}_m &= (a_{m1}, \ldots, a_{mn}), \\ \mathbf{b} &= (b_1, \ldots, b_n), \end{aligned}$$

l'égalité vectorielle

$$x_1\mathbf{a}_1 + \ldots + x_m\mathbf{a}_m = \mathbf{b} \tag{5}$$

est équivalente à n égalités numériques

$$\begin{aligned} a_{11}x_1 + \ldots + a_{m1}x_m &= b_1, \\ &\ldots\ldots\ldots \\ a_{1n}x_1 + \ldots + a_{mn}x_m &= b_n. \end{aligned} \tag{6}$$

Les relations (6) constituent un système de n équations linéaires non homogènes à m inconnues. Ce système est *compatible*, i.e. il possède au moins une solution $x_1, \ldots, x_m$ si et seulement si l'on a (5), i.e. si le vecteur $\mathbf{b}$ s'exprime linéairement en fonction de $\mathbf{a}_1, \ldots, \mathbf{a}_m$.

D'autre part, le rang de l'ensemble $\mathbf{a}_1, \ldots, \mathbf{a}_m$ est égal, par le théorème 1, au rang de la *matrice des coefficients*

$$\begin{pmatrix} a_{11} & \ldots & a_{m1} \\ \ldots & \ldots & \ldots \\ a_{1n} & \ldots & a_{mn} \end{pmatrix} \tag{7}$$

du système (6), et le rang de l'ensemble $\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}$ au rang de la *matrice allongée des coefficients*

$$(8) \qquad \begin{pmatrix} a_{11} & \ldots & a_{m1} & b_1 \\ \cdot & \cdot & \cdot & \cdot \\ a_{1n} & \ldots & a_{mn} & b_n \end{pmatrix}$$

obtenue de (7) par adjonction de la colonne des seconds membres.

On a donc le

Théorème 2 (Kronecker-Capelli). *Le système d'équations linéaires* (6) *est compatible si et seulement si la matrice de ses coefficients* (7) *et la matrice allongée* (8) *ont même rang.*

* * *

Supposons qu'il y a compatibilité. Comment trouver toutes les solutions de (6)?

Soit r le rang de la matrice (7). On permute les équations, on change (si besoin est) les notations des inconnues et on estime sans restreindre la généralité que

$$(9) \qquad \Delta = \begin{vmatrix} a_{11} & \ldots & a_{r1} \\ \cdot & \cdot & \cdot \\ a_{1r} & \ldots & a_{rr} \end{vmatrix} \neq 0.$$

Le système (6) étant compatible par hypothèse, le théorème de Kronecker-Capelli veut que le rang de la matrice (8) soit r. Par suite de la condition (9), cela signifie que r premières lignes de (8) (i.e. r premières équations (6)) sont linéairement indépendantes et que toute autre ligne de (8) (toute autre équation (6)) en est combinaison linéaire. Aussi le système (6) est équivalent au système

$$(10) \qquad \begin{aligned} a_{11}x_1 + \ldots + a_{r1}x_r + \ldots + a_{m1}x_m &= b_1, \\ \cdot \quad \cdot \quad \cdot \quad

de Cramer du cours d'Algèbre nous permettent donc de déterminer de façon unique $x_1, \ldots, x_r$. Il est clair qu'avec ce procédé, on obtient *toutes* les solutions de (10) (i.e. de (6)).

Dans la pratique, il est certes inutile de permuter les équations et de changer les notations des inconnues. La résolution d'un système linéaire (6) arbitraire s'effectue donc par étapes.

Etape première. On calcule les mineurs de la matrice des coefficients (7) et on établit son rang r tout en trouvant au moins un mineur Δ d'ordre r non nul.

Etape 2. On borde le mineur trouvé dans (8) et on constate que cette matrice est de rang r. (Si le rang est $r + 1$, il y a incompatibilité de (6).) A cette étape, il suffit visiblement de calculer $n - r$ mineurs d'ordre $r + 1$.

Etape 3. Dans le mineur Δ figurent les coefficients précédant r inconnues de r équations. On chasse les autres équations, on fait prendre aux $n - r$ inconnues restantes des valeurs quelconques et on obtient donc un système de r équations à r inconnues dont le déterminant est non nul. On résout le système moyennant les formules de Cramer et on trouve les valeurs des r inconnues.

Les valeurs des $x_1, \ldots, x_m$ ainsi obtenues sont solutions du système (6), et toute solution de (6) s'obtient de cette façon.

LEÇON 3

Somme directe de sous-espaces vectoriels. — Décomposition d'un espace vectoriel en somme directe de sous-espaces. — Espace vectoriel quotient. — Homomorphismes d'espaces vectoriels. — Somme directe d'espaces vectoriels.

Soit $\mathcal{P}$ et $\mathcal{Q}$ deux sous-espaces de l'espace vectoriel $\mathcal{V}$. On rappelle que la somme $\mathcal{P} + \mathcal{Q}$ est formée de tous les vecteurs de la forme $\mathbf{x} + \mathbf{y}$, où $\mathbf{x} \in \mathcal{P}$, $\mathbf{y} \in \mathcal{Q}$.

Définition 1. On dit que le sous-espace $\mathcal{P} + \mathcal{Q}$ est la *somme directe* des sous-espaces $\mathcal{P}$ et $\mathcal{Q}$ si chaque vecteur de $\mathcal{P} + \mathcal{Q}$ est représentable de façon unique sous la forme $\mathbf{x} + \mathbf{y}$, $\mathbf{x} \in \mathcal{P}$, $\mathbf{y} \in \mathcal{Q}$.

La somme directe de $\mathcal{P}$ et $\mathcal{Q}$ se note $\mathcal{P} \oplus \mathcal{Q}$ ou $\mathcal{P} \dot{+} \mathcal{Q}$.

Proposition 1. *Le sous-espace $\mathcal{P} + \mathcal{Q}$ est somme directe des sous-espaces $\mathcal{P}$ et $\mathcal{Q}$ si et seulement si $\mathcal{P} \cap \mathcal{Q} = \mathbf{0}$.*

Démonstration. Si $\mathbf{x} + \mathbf{y} = \mathbf{x}_1 + \mathbf{y}_1$, avec $\mathbf{x}, \mathbf{x}_1 \in \mathcal{P}$ et $\mathbf{y}, \mathbf{y}_1 \in \mathcal{Q}$, alors le vecteur $\mathbf{x} - \mathbf{x}_1 = \mathbf{y}_1 - \mathbf{y}$ est dans $\mathcal{P} \cap \mathcal{Q}$. Aussi $\mathbf{x} = \mathbf{x}_1$ et $\mathbf{y} = \mathbf{y}_1$ si $\mathcal{P} \cap \mathcal{Q} = \mathbf{0}$, i.e. il y a unicité de la représentation $\mathbf{x} + \mathbf{y}$, $\mathbf{x} \in \mathcal{P}$, $\mathbf{y} \in \mathcal{Q}$, pour chaque vecteur de $\mathcal{P} + \mathcal{Q}$. Soit maintenant $\mathcal{P} \cap \mathcal{Q} \neq \mathbf{0}$ et $\mathbf{a} \in \mathcal{P} \cap \mathcal{Q}$, $\mathbf{a} \neq \mathbf{0}$. On a pour tout $\mathbf{x} \in \mathcal{P}$ et tout $\mathbf{y} \in \mathcal{Q}$

$$\mathbf{x} + \mathbf{y} = (\mathbf{x} + \mathbf{a}) + (\mathbf{y} - \mathbf{a}),$$

avec $\mathbf{x} + \mathbf{a} \in \mathcal{P}$ et $\mathbf{y} - \mathbf{a} \in \mathcal{Q}$. Cette égalité montre que les vecteurs de $\mathcal{P} + \mathcal{Q}$ admettent d'autres représentations à part $\mathbf{x} + \mathbf{y}$, $\mathbf{x} \in \mathcal{P}$, $\mathbf{y} \in \mathcal{Q}$. □

La notion de somme directe de sous-espaces en nombre quelconque a certes un sens. Ainsi, la somme $\mathcal{P} + \mathcal{Q} + \mathcal{R}$ de trois sous-espaces est dite directe si chaque vecteur de $\mathcal{P} + \mathcal{Q} + \mathcal{R}$ se met de façon unique sous la forme $\mathbf{x} + \mathbf{y} + \mathbf{z}$, $\mathbf{x} \in \mathcal{P}$, $\mathbf{y} \in \mathcal{Q}$, $\mathbf{z} \in \mathcal{R}$. On croirait, par analogie avec la proposition 1, qu'une condition nécessaire et suffisante en est que $\mathcal{P}$, $\mathcal{Q}$ et $\mathcal{R}$ soient disjoints deux à deux. Or, *il n'en est rien*. Etant donnés, par exemple, deux vecteurs non colinéaires quelconques $\mathbf{a}$ et $\mathbf{b}$, les sous-espaces $\mathcal{P} = [\mathbf{a}]$, $\mathcal{Q} = [\mathbf{b}]$, $\mathcal{R} = [\mathbf{a} + \mathbf{b}]$ sont disjoints deux à deux, mais leur somme $\mathcal{P} + \mathcal{Q} + \mathcal{R} = [\mathbf{a}, \mathbf{b}]$ n'est pas somme directe.

La condition pour que la somme $\mathcal{P} + \mathcal{Q} + \mathcal{R}$ soit directe est donnée par la

Proposition 2. *Pour que la somme $\mathcal{P} + \mathcal{Q} + \mathcal{R}$ de trois sous-espaces soit directe, il faut et il suffit que chaque sous-espace soit disjoint de la somme de deux autres:*

$$(1) \qquad \mathcal{P} \cap (\mathcal{Q} + \mathcal{R}) = 0, \quad \mathcal{Q} \cap (\mathcal{P} + \mathcal{R}) = 0, \quad \mathcal{R} \cap (\mathcal{P} + \mathcal{Q}) = 0.$$

Démonstration. Si $\mathbf{x} + \mathbf{y} + \mathbf{z} = \mathbf{x}_1 + \mathbf{y}_1 + \mathbf{z}_1$, avec $\mathbf{x}, \mathbf{x}_1 \in \mathcal{P}$, $\mathbf{y}, \mathbf{y}_1 \in \mathcal{Q}$, $\mathbf{z}, \mathbf{z}_1 \in \mathcal{R}$, alors $\mathbf{x} - \mathbf{x}_1 = (\mathbf{y}_1 - \mathbf{y}) + (\mathbf{z}_1 - \mathbf{z}) \in \mathcal{P} \cap \cap (\mathcal{Q} + \mathcal{R})$. Donc $\mathcal{P} \cap (\mathcal{Q} + \mathcal{R}) \neq 0$ si $\mathbf{x}_1 \neq \mathbf{x}$. De même, $\mathcal{Q} \cap (\mathcal{P} + \mathcal{R}) \neq 0$ si $\mathbf{y}_1 \neq \mathbf{y}$ et $\mathcal{R} \cap (\mathcal{P} + \mathcal{Q}) \neq 0$ pour $\mathbf{z}_1 \neq \mathbf{z}$. Ainsi, si la somme $\mathcal{P} + \mathcal{Q} + \mathcal{R}$ n'est pas directe, toutes les conditions (1) ne sont pas remplies. Inversement, soit, par exemple $\mathcal{P} \cap (\mathcal{Q} + \mathcal{R}) \neq 0$ et $\mathbf{a} \in \mathcal{P} \cap (\mathcal{Q} + \mathcal{R})$, $\mathbf{a} \neq 0$. N'importe quels $\mathbf{x} \in \mathcal{P}$, $\mathbf{y} \in \mathcal{Q}$ et $\mathbf{z} \in \mathcal{R}$ vérifient alors l'égalité

$$\mathbf{x} + \mathbf{y} + \mathbf{z} = (\mathbf{x} - \mathbf{a}) + (\mathbf{y} + \mathbf{b}) + (\mathbf{z} + \mathbf{c}),$$

$\mathbf{b} \in \mathcal{Q}$, $\mathbf{c} \in \mathcal{R}$ étant des vecteurs tels que $\mathbf{a} = \mathbf{b} + \mathbf{c}$, si bien que la somme $\mathcal{P} + \mathcal{Q} + \mathcal{R}$ n'est pas directe. □

On démontre certes une proposition analogue pour plus de trois sous-espaces.

* * *

Le cas le plus intéressant est celui où $\mathcal{P} \oplus \mathcal{Q} = \mathcal{V}$. On dit que l'espace $\mathcal{V}$ *est décomposé en somme directe* des sous-espaces $\mathcal{P}$ et $\mathcal{Q}$.

Voyons certaines propriétés de $\mathcal{P}$ et $\mathcal{Q}$:

1° Tout vecteur de $\mathcal{V}$ s'écrit comme $\mathbf{x} + \mathbf{y}$, $\mathbf{x} \in \mathcal{P}$, $\mathbf{y} \in \mathcal{Q}$, i.e. $\mathcal{V} = \mathcal{P} + \mathcal{Q}$.

2° Les sous-espaces $\mathcal{P}$ et $\mathcal{Q}$ sont disjoints, i.e. $\mathcal{P} \cap \mathcal{Q} = 0$.

3° La somme des dimensions des sous-espaces $\mathcal{P}$ et $\mathcal{Q}$ est égale à la dimension de $\mathcal{V}$:

$$\dim \mathcal{P} + \dim \mathcal{Q} = \dim \mathcal{V}.$$

Proposition 3. *Deux quelconques des propriétés* 1°, 2°, 3° *entraînent la troisième.*

Démonstration. S'il y a 1° et 2°, le théorème 1 de la leçon 1 (théorème sur la dimension de la somme) implique

$$\dim \mathcal{V} = \dim (\mathcal{P} + \mathcal{Q}) = \\ = \dim \mathcal{P} + \dim \mathcal{Q} + \dim (\mathcal{P} \cap \mathcal{Q}) = \dim \mathcal{P} + \dim \mathcal{Q}$$

Si l'on est dans les cas 1° et 3°, on a par le même théorème

$$\dim (\mathcal{P} \cap \mathcal{Q}) = \dim (\mathcal{P} + \mathcal{Q}) - \dim \mathcal{P} - \dim \mathcal{Q} = \\ = \dim \mathcal{V} - \dim \mathcal{P} - \dim \mathcal{Q} = 0,$$

donc $\mathcal{P} \cap \mathcal{Q} = 0$.

Soit enfin les propriétés 2° et 3°. Le théorème mentionné donne

$$\dim (\mathcal{P} + \mathcal{Q}) = \dim \mathcal{P} + \dim \mathcal{Q} = \dim \mathcal{V},$$

si bien que $\mathcal{P} + \mathcal{Q} = \mathcal{V}$. □

D'après la proposition 1, les propriétés 1° et 2° signifient que $\mathcal{V} = \mathcal{P} \oplus \mathcal{Q}$. Ainsi on a la

Conséquence. *L'égalité $\mathcal{V} = \mathcal{P} \oplus \mathcal{Q}$ a lieu si et seulement si les sous-espaces possèdent deux quelconques des propriétés* 1°, 2°, 3° (*donc les trois propriétés*).

Définition 2. Si $\mathcal{V} = \mathcal{P} \oplus \mathcal{Q}$, les sous-espaces $\mathcal{P}$ et $\mathcal{Q}$ sont dits *supplémentaires.*

Proposition 4. *Etant donnés $\mathcal{P}$ et $\mathcal{Q}$ supplémentaires, toute base* $\mathbf{e}_1, \ldots, \mathbf{e}_p$ *du sous-espace $\mathcal{P}$ et toute base* $\mathbf{e}_{p+1}, \ldots, \mathbf{e}_n$ *du sous-espace $\mathcal{Q}$, les vecteurs*

$$\mathbf{e}_1, \ldots, \mathbf{e}_p, \mathbf{e}_{p+1}, \ldots, \mathbf{e}_n$$

forment une base de l'espace $\mathcal{V}$.

Inversement, si l'on décompose une base quelconque $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *de l'espace $\mathcal{V}$ en deux sous-familles* $\mathbf{e}_1, \ldots, \mathbf{e}_p$ *et* $\mathbf{e}_{p+1}, \ldots, \mathbf{e}_n$, *alors les sous-espaces* $\mathcal{P} = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$ *et* $\mathcal{Q} = [\mathbf{e}_{p+1}, \ldots, \mathbf{e}_n]$ *sont supplémentaires.*

Démonstration. Les vecteurs $\mathbf{e}_1, \ldots, \mathbf{e}_p, \mathbf{e}_{p+1}, \ldots, \mathbf{e}_n$ de la première affirmation constituent une famille génératrice de $n = p + q$ vecteurs, i.e. c'est une base. Dans la seconde affirmation, les sous-espaces $\mathcal{P}$ et $\mathcal{Q}$ jouissent des propriétés 1° et 3°, donc $\mathcal{V} = \mathcal{P} \oplus \mathcal{Q}$. □

Conséquence. *Pour tout sous-espace $\mathcal{P} \subset \mathcal{V}$ il existe un supplémentaire $\mathcal{Q}$.*

Démonstration. Soit $\mathbf{e}_1, \ldots, \mathbf{e}_p$ une base quelconque du sous-espace $\mathcal{P}$. Complétons-la par des vecteurs $\mathbf{e}_{p+1}, \ldots, \mathbf{e}_n$ de façon à former une base de l'espace $\mathcal{V}$ tout entier. Alors le sous-espace $\mathcal{Q} = [\mathbf{e}_{p+1}, \ldots, \mathbf{e}_n]$ est supplémentaire de $\mathcal{P}$. □

On voit que la construction du sous-espace supplémentaire $\mathcal{Q}$ comporte un grand arbitraire qui peut d'ailleurs être atténué (sinon supprimé) par un certain procédé.

* * *

Soit $\mathcal{P}$ un sous-espace quelconque de l'espace vectoriel $\mathcal{V}$.

Définition 3. Deux vecteurs $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ sont dits *congrus modulo* $\mathcal{P}$ si $\mathbf{x} - \mathbf{y} \in \mathcal{P}$. La notation correspondante est

$$\mathbf{x} \equiv \mathbf{y} \bmod \mathcal{P}.$$

La relation de congruence est évidemment une relation d'équivalence. Les classes de vecteurs congrus modulo $\mathcal{P}$ s'appellent *classes de $\mathcal{V}$ suivant le sous-espace $\mathcal{P}$*. Il est clair que la classe contenant $\mathbf{x}$

a pour éléments tous les vecteurs de la forme $\mathbf{x} + \mathbf{a}$, $\mathbf{a} \in \mathcal{P}$. Cette classe sera notée $\mathbf{x} + \mathcal{P}$ (ou encore $\mathbf{x} \bmod \mathcal{P}$).

On constate sans peine qu'on peut additionner les congruences et les multiplier par des nombres, i.e. si

$$\mathbf{x} \equiv \mathbf{y} \bmod \mathcal{P} \quad \text{et} \quad \mathbf{x}_1 \equiv \mathbf{y}_1 \bmod \mathcal{P},$$

alors

$$\mathbf{x} + \mathbf{x}_1 \equiv \mathbf{y} + \mathbf{y}_1 \bmod \mathcal{P}$$

et

$$k\mathbf{x} \equiv k\mathbf{y} \bmod \mathcal{P}$$

quel que soit $k \in \mathrm{K}$. En effet, si $\mathbf{x} - \mathbf{y} \in \mathcal{P}$ et $\mathbf{x}_1 - \mathbf{y}_1 \in \mathcal{P}$, alors $(\mathbf{x} + \mathbf{x}_1) - (\mathbf{y} + \mathbf{y}_1) = (\mathbf{x} - \mathbf{y}) + (\mathbf{x}_1 - \mathbf{y}_1) \in \mathcal{P}$, et, de même, $k\mathbf{x} - k\mathbf{y} = k(\mathbf{x} - \mathbf{y}) \in \mathcal{P}$.

S'agissant des classes suivant le sous-espace $\mathcal{P}$, cela signifie que les formules

$$(\mathbf{x} + \mathcal{P}) + (\mathbf{y} + \mathcal{P}) = (\mathbf{x} + \mathbf{y}) + \mathcal{P} \tag{2}$$

et

$$k(\mathbf{x} + \mathcal{P}) = k\mathbf{x} + \mathcal{P} \tag{3}$$

définissent bien leur somme et leur produit par un nombre.

On vérifie de suite que ces opérations respectent les axiomes des espaces vectoriels. Ainsi, *l'ensemble des classes de $\mathcal{V}$ suivant $\mathcal{P}$ est un espace vectoriel* pour les opérations (2) et (3). □

Définition 4. Cet espace s'appelle l'*espace vectoriel quotient de $\mathcal{V}$ par $\mathcal{P}$ et* se note $\mathcal{V}/\mathcal{P}$.

Le cas analogue de groupes et d'anneaux a été étudié en détail dans le cours d'Algèbre (premier semestre).

Proposition 5. *Chaque sous-espace $\mathcal{Q}$ supplémentaire de sous-espace $\mathcal{P}$ est isomorphe à l'espace quotient $\mathcal{V}/\mathcal{P}$.*

Démonstration. Soit l'application $\varphi : \mathcal{Q} \to \mathcal{V}/\mathcal{P}$ définie par la formule

$$\varphi(\mathbf{x}) = \mathbf{x} + \mathcal{P}, \quad \mathbf{x} \in \mathcal{Q}.$$

Si $\varphi(\mathbf{x}) = \varphi(\mathbf{x}_1)$, i.e. si $\mathbf{x} + \mathcal{P} = \mathbf{x}_1 + \mathcal{P}$, alors $\mathbf{x} - \mathbf{x}_1 \in \mathcal{P}$, donc $\mathbf{x} = \mathbf{x}_1$. D'autre part, tout vecteur $\mathbf{z} \in \mathcal{V}$ s'écrit $\mathbf{x} + \mathbf{y}$, avec $\mathbf{x} \in \mathcal{Q}$, $\mathbf{y} \in \mathcal{P}$, donc $\mathbf{z} + \mathcal{P} = \mathbf{x} + \mathcal{P}$. Cela démontre la nature bijective de φ. Puisque l'application φ conserve évidemment les sommes et les produits par des nombres, c'est un isomorphisme. □

Géométriquement, la proposition 5 part du fait que chaque classe suivant $\mathcal{P}$ et $\mathcal{Q}$ ont exactement un vecteur commun.

La proposition 5 signifie qu'on considère, au lieu des supplémentaires $\mathcal{Q}$, l'espace quotient $\mathcal{V}/\mathcal{P}$ qu'on construit sans se permettre aucun arbitraire.

Elle entraîne

(4) $$\dim \mathscr{V}/\mathscr{P} = \dim \mathscr{V} - \dim \mathscr{P}.$$

En effet, $\dim \mathscr{V}/\mathscr{P} = \dim \mathscr{Q} = \dim \mathscr{V} - \dim \mathscr{P}$. □

* * *

Soient $\mathscr{V}$ et $\mathscr{W}$ deux espaces vectoriels.

Définition 5. L'application

$$\varphi : \mathscr{V} \to \mathscr{W}$$

s'appelle *application linéaire* ou *homomorphisme* (ou encore *morphisme*) d'espaces vectoriels si elle conserve les opérations linéaires, i.e. si

$$\varphi(\mathbf{x} + \mathbf{y}) = \varphi(\mathbf{x}) + \varphi(\mathbf{y})$$

et

$$\varphi(k\mathbf{x}) = k\varphi(\mathbf{x})$$

pour $\mathbf{x}, \mathbf{y} \in \mathscr{V}$ quelconques et tout $k \in \mathbb{K}$.

Ainsi, un isomorphisme est un homomorphisme bijectif.

Définition 6. L'ensemble des vecteurs $\mathbf{x} \in \mathscr{V}$ qui deviennent par un homomorphisme φ l'élément $\mathbf{0}$ de l'espace $\mathscr{W}$ s'appelle le *noyau* de l'homomorphisme φ et se note $\operatorname{Ker} \varphi$. Ainsi,

$$\operatorname{Ker} \varphi = \{\mathbf{x} \in \mathscr{V};\ \varphi(\mathbf{x}) = \mathbf{0}\}.$$

Définition 7. L'ensemble des vecteurs de $\mathscr{W}$ de la forme $\varphi(\mathbf{x})$, $\mathbf{x} \in \mathscr{V}$, s'appelle l'*image* de l'homomorphisme φ et se note $\operatorname{Im} \varphi$:

$$\operatorname{Im} \varphi = \{\mathbf{y} \in \mathscr{W};\ \mathbf{y} = \varphi(\mathbf{x})\}.$$

L'image de φ est également noté $\varphi(\mathscr{V})$ (on parle alors de l'*image de l'espace* $\mathscr{V}$ par l'homomorphisme φ).

Evidemment, *les ensembles* $\operatorname{Ker} \varphi$ *et* $\operatorname{Im} \varphi$ *sont des sous-espaces* (de $\mathscr{V}$ et $\mathscr{W}$ respectivement).

L'espace quotient $\mathscr{W}/\operatorname{Im} \varphi$ est appelé *conoyau* de l'homomorphisme φ et noté $\operatorname{Coker} \varphi$.

Etant donné un homomorphisme φ, si φ est une application injective, i.e. si $\varphi(\mathbf{x}) = \varphi(\mathbf{x}_1)$ pour $\mathbf{x} \neq \mathbf{x}_1$, on dit que c'est un *monomorphisme*.

Un homomorphisme φ s'appelle *épimorphisme* s'il applique $\mathscr{V}$ sur $\mathscr{W}$, i.e. si l'on associe à chaque vecteur $\mathbf{y} \in \mathscr{W}$ un vecteur $\mathbf{x} \in \mathscr{V}$ tel que $\mathbf{y} = \varphi(x)$.

Ainsi, une condition nécessaire et suffisante pour qu'un homomorphisme φ soit un isomorphisme est qu'il soit simultanément un monomorphisme et un épimorphisme.

Par définition, *un homomorphisme* φ *est un épimorphisme si et seulement si* $\operatorname{Im} \varphi = \mathscr{W}$, *i.e. si* $\operatorname{Coker} \varphi = \mathbf{0}$. □

De même, on établit sans peine qu'*un homomorphisme* φ *est un monomorphisme si et seulement si* $\operatorname{Ker}\varphi = 0$. En effet, si $\varphi(x) = \varphi(x_1)$, alors $\varphi(x - x_1) = 0$, donc $x - x_1 \in \operatorname{Ker}\varphi$. Si $\operatorname{Ker}\varphi = 0$, on a donc $x = x_1$. Inversement, si l'égalité $\varphi(x) = \varphi(x_1)$ entraîne $x = x_1$, alors on a en particulier $\varphi(x) = 0$ si et seulement si $x = 0$. Par conséquent, $\operatorname{Ker}\varphi = 0$. □

Si $\operatorname{Ker}\varphi = 0$, l'application φ est évidemment un isomorphisme de l'espace $\mathcal{V}$ sur le sous-espace $\operatorname{Im}\varphi \subset \mathcal{W}$. Aussi $\dim \operatorname{Im}\varphi = \dim \mathcal{V}$. Il en résulte que *si* $\operatorname{Ker}\varphi = 0$ *et* $\dim \mathcal{V} = \dim \mathcal{W}$, *alors l'homomorphisme* φ *est un isomorphisme*. En effet, on a dans ce cas $\dim \operatorname{Im}\varphi = \dim \mathcal{W}$ et, partant, $\operatorname{Im}\varphi = \mathcal{W}$. □

Quand $\operatorname{Ker}\varphi \neq 0$, il y a intérêt à introduire l'espace quotient

$$\mathcal{V}/\operatorname{Ker}\varphi$$

dont on dit des fois que c'est la *coimage* de l'homomorphisme φ. La formule

$$\varphi'(x + \mathcal{P}) = \varphi(x), \quad x \in \mathcal{V}.$$

définit visiblement bien un homomorphisme

$$\varphi' : \mathcal{V}/\operatorname{Ker}\varphi \to \mathcal{W}$$

appelé *homomorphisme induit*, et l'homomorphisme φ' est, on le voit facilement, un isomorphisme de l'espace quotient $\mathcal{V}/\operatorname{Ker}\varphi$ sur le sous-espace $\operatorname{Im}\varphi$.

En particulier, on montre que *pour tout épimorphisme* $\varphi : \mathcal{V} \to \mathcal{W}$, *l'espace* $\mathcal{W}$ *est isomorphe à l'espace quotient* $\mathcal{V}/\operatorname{Ker}\varphi$. □

Comme $\dim \mathcal{V}/\operatorname{Ker}\varphi = \dim \mathcal{V} - \dim \operatorname{Ker}\varphi$, on énonce de plus un autre résultat : *pour tout homomorphisme* $\varphi : \mathcal{V} \to \mathcal{W}$ *on a la formule*

$$\dim \operatorname{Ker}\varphi + \dim \operatorname{Im}\varphi = \dim \mathcal{V}. \tag{5}$$

La formule (5) exceptée, toutes les affirmations ci-dessus sont très générales. Elles sont valables pour les groupes et les anneaux quelconques (voir le cours d'Algèbre du premier semestre).

* * *

Revenons aux sommes directes.

Soit $\mathcal{P}$ et $\mathcal{Q}$ deux espaces vectoriels arbitraires (sur un même corps de base K) et soit $\mathcal{V}$ l'ensemble de tous les couples (x, y), $x \in \mathcal{P}$, $y \in \mathcal{Q}$. On pose

$$(x, y) + (x_1, y_1) = (x + x_1, y + y_1)$$

et

$$k(x, y) = (kx, ky),$$

et on structure évidemment $\mathcal{V}$ en espace vectoriel.

Définition 8. L'espace $\mathcal{V}$ ainsi construit s'appelle *somme directe* des espaces $\mathcal{P}$ et $\mathcal{Q}$ (ou encore *somme directe externe* pour la distinguer d'avec la somme directe « interne » du début de la leçon, où l'espace $\mathcal{V}$ a été donné à l'avance, et $\mathcal{P}$ et $\mathcal{Q}$ en ont été des sous-espaces).

Cette terminologie tient à ce que les vecteurs $(\mathbf{x}, 0) \in \mathcal{V}$, où $\mathbf{x} \in \mathcal{P}$, constituent un sous-espace $\hat{\mathcal{P}}$ isomorphe à l'espace $\mathcal{P}$, et les vecteurs $(0, \mathbf{y}) \in \mathcal{V}$, où $\mathbf{y} \in \mathcal{Q}$, forment un sous-espace $\hat{\mathcal{Q}}$ isomorphe à l'espace $\mathcal{Q}$. Les sous-espaces $\hat{\mathcal{P}}$ et $\hat{\mathcal{Q}}$ sont de plus disjoints (i.e. leur seul élément commun est le vecteur nul $(0, 0)$) et remplissent ensemble l'espace $\mathcal{V}$ tout entier (car $(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, 0) + (0, \mathbf{y})$). Ainsi, $\mathcal{V} = \hat{\mathcal{P}} \oplus \hat{\mathcal{Q}}$.

L'usage veut qu'on identifie $\hat{\mathcal{P}}$ et $\hat{\mathcal{Q}}$ avec $\mathcal{P}$ et $\mathcal{Q}$ respectivement. Ce fait est noté $\mathcal{V} = \mathcal{P} \oplus \mathcal{Q}$ ou $\mathcal{V} = \mathcal{P} \dot{+} \mathcal{Q}$, ce qui n'implique aucune ambiguïté.

S'agissant des groupes, la somme directe externe est construite dans le cours d'Algèbre (premier semestre). Nous l'avons en fait utilisée dans I pour obtenir des complexifiés.

Dans la leçon suivante, nous considérerons des constructions plus naturelles à la théorie des espaces vectoriels.

LEÇON 4

Espace dual. — Espaces en dualité. — Bidual. — Transformation de la base duale et des coordonnées des covecteurs. — Annulateurs. — Espace des solutions d'un système d'équations linéaires homogènes.

Soit $\mathscr{V}$ un espace vectoriel quelconque sur $\mathbb{K}$.

Définition 1. Une fonction $\xi : \mathscr{V} \to \mathbb{K}$ s'appelle *fonctionnelle linéaire* si elle est un homomorphisme d'espaces vectoriels, i.e. si

$$\xi(\mathbf{x} + \mathbf{y}) = \xi(\mathbf{x}) + \xi(\mathbf{y})$$

et

$$\xi(k\mathbf{x}) = k\xi(\mathbf{x})$$

pour $\mathbf{x}, \mathbf{y} \in \mathscr{V}$ quelconques et tout $k \in \mathbb{K}$. Les fonctionnelles linéaires sont également appelées *covecteurs* de l'espace $\mathscr{V}$.

Il est immédiat de vérifier que la somme $\xi + \eta$ de deux fonctionnelles linéaires ξ et η (qu'on définit par la formule $(\xi + \eta)(\mathbf{x}) = \xi(\mathbf{x}) + \eta(\mathbf{x})$) et le produit $k\xi$ d'une fonctionnelle linéaire ξ par un nombre quelconque k (qu'on définit par la formule $(k\xi)(\mathbf{x}) = k\xi(\mathbf{x})$) sont des fonctionnelles linéaires. Autrement dit, l'ensemble de toutes les fonctionnelles linéaires constitue un sous-espace de l'espace de toutes les fonctions sur $\mathscr{V}$, i.e. c'est un espace vectoriel. On désigne cet espace vectoriel par $\mathsf{T}_1(\mathscr{V})$ ou $\mathscr{V}'$.

Définition 2. L'espace vectoriel $\mathscr{V}'$ est dit *dual* de l'espace $\mathscr{V}$.

Soit $\mathbf{e}_1, \ldots, \mathbf{e}_n$ une base arbitraire de $\mathscr{V}$.

Proposition 1. *La valeur $\xi(\mathbf{x})$ d'une fonctionnelle linéaire quelconque ξ sur le vecteur $\mathbf{x} = x^1\mathbf{e}_1 + \ldots + x^n\mathbf{e}_n$ est donnée par la formule*

$$\xi(\mathbf{x}) = \xi_1 x^1 + \ldots + \xi_n x^n, \tag{1}$$

où

$$\xi_1 = \xi(\mathbf{e}_1), \ldots, \xi_n = \xi(\mathbf{e}_n). \tag{2}$$

Etant donnés des nombres $\xi_1, \ldots, \xi_n \in \mathbb{K}$ quelconques, la formule (1) *définit de façon unique une fonctionnelle linéaire $\xi \in \mathscr{V}'$ qui vérifie l'égalité* (2).

Démonstration. La formule (1) résulte de suite de la linéarité

$$\xi(\mathbf{x}) = \xi(x^1\mathbf{e}_1 + \ldots + x^n\mathbf{e}_n) = \\ = x^1\xi(\mathbf{e}_1) + \ldots + x^n\xi(\mathbf{e}_n) = \xi_1 x^1 + \ldots + \xi_n x^n.$$

Inversement, soit ξ définie par (1). On a

$$\xi(\mathbf{x}+\mathbf{y}) = \xi_1(x^1+y^1) + \ldots + \xi_n(x^n+y^n) = \\ = \xi_1 x^1 + \ldots + \xi_n x^n + \xi_1 y^1 + \ldots + \xi_n y^n = \xi(\mathbf{x}) + \xi(\mathbf{y})$$

et

$$\xi(k\mathbf{x}) = \xi_1(kx^1) + \ldots + \xi_n(kx^n) = \\ = k(\xi_1 x^1 + \ldots + \xi_n x^n) = k\xi(\mathbf{x})$$

quels que soient $\mathbf{x}, \mathbf{y} \in \mathscr{V}$ et $k \in \mathbb{K}$. De plus, $\xi(\mathbf{e}_i) = \xi_1 \cdot 0 + \ldots + \xi_i \cdot 1 + \ldots + \xi_n \cdot 0 = \xi_i$. □

La proposition 1 entraîne la propriété de

$$\mathbf{e}^i(\mathbf{e}_j) = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j, \end{cases} \quad i, j = 1, \ldots, n,$$

de définir univoquement n fonctionnelles linéaires

$$\mathbf{e}^1, \ldots, \mathbf{e}^n. \tag{3}$$

Il est clair que

$$\mathbf{e}^i(\mathbf{x}) = x^i, \quad i = 1, \ldots, n,$$

pour tout $\mathbf{x} \in \mathscr{V}$.

Proposition 2. *Les fonctionnelles* (3) *forment une base de l'espace* $\mathscr{V}'$. *Les coordonnées d'une fonctionnelle quelconque* ξ *dans cette base sont les coefficients* (2) *de sa représentation* (1):

$$\xi = \xi_1 \mathbf{e}^1 + \ldots + \xi_n \mathbf{e}^n. \tag{4}$$

Démonstration. Pour tout vecteur $\mathbf{x} = x^1\mathbf{e}_1 + \ldots + x^n\mathbf{e}_n$ et n'importe quels nombres $\xi_1, \ldots, \xi_n \in \mathbb{K}$, on a

$$(\xi_1\mathbf{e}^1 + \ldots + \xi_n\mathbf{e}^n)(\mathbf{x}) = \xi_1\mathbf{e}^1(\mathbf{x}) + \ldots + \xi_n\mathbf{e}^n(\mathbf{x}) = \\ = \xi_1 x^1 + \ldots + \xi_n x^n.$$

Si $\xi_1, \ldots, \xi_n$ sont les coefficients (2) de la fonctionnelle ξ, alors on a donc $(\xi_1\mathbf{e}^1 + \ldots + \xi_n\mathbf{e}^n)(\mathbf{x}) = \xi(\mathbf{x})$ pour tout $\mathbf{x} \in \mathscr{V}$. La formule (4) se trouve démontrée, ainsi que la propriété, pour la famille $\mathbf{e}^1, \ldots, \mathbf{e}^n$, d'être génératrice dans $\mathscr{V}'$.

D'autre part, si

$$\xi_1\mathbf{e}^1 + \ldots + \xi_n\mathbf{e}^n = 0,$$

alors on a pour tout $i = 1, \ldots, n$:

$$\xi_i = (\xi_1\mathbf{e}^1 + \ldots + \xi_n\mathbf{e}^n)(\mathbf{e}_i) = 0.$$

Par conséquent, il y a indépendance linéaire de $\mathbf{e}^1, \ldots, \mathbf{e}^n$ qui est donc une base. □

Conséquence.

$$\dim \mathscr{V}' = \dim \mathscr{V}.$$

La base $\mathbf{e}^1, \ldots, \mathbf{e}^n$ est dite *duale* de la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

Avec les notations d'Einstein, la formule (1) s'écrit

$$\xi(\mathbf{x}) = \xi_i x^i,$$

et la formule (4)

$$\xi = \xi_i \mathbf{e}^i.$$

* * *

Soient $\mathscr{V}$ et $\mathscr{W}$ deux espaces vectoriels sur $\mathbb{K}$. On suppose qu'il existe, pour tout couple de vecteurs $\mathbf{x}$, $\mathbf{y}$, où $\mathbf{x} \in \mathscr{V}$, $\mathbf{y} \in \mathscr{W}$, un nombre $\langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{K}$ tel que

a) étant donné un vecteur $\mathbf{y} \in \mathscr{W}$ fixe quelconque, la fonction $\mathbf{x} \mapsto \langle \mathbf{x}, \mathbf{y} \rangle$ est une fonctionnelle linéaire sur $\mathscr{V}$, i. e.

$$\langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{y} \rangle = \langle \mathbf{x}_1, \mathbf{y} \rangle + \langle \mathbf{x}_2, \mathbf{y} \rangle,$$
$$\langle k\mathbf{x}, \mathbf{y} \rangle = k \langle \mathbf{x}, \mathbf{y} \rangle$$

pour n'importe quels $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x} \in \mathscr{V}$ et tout $k \in \mathbb{K}$;

b) étant donné un vecteur $\mathbf{x} \in \mathscr{V}$ fixe quelconque, la fonction $\mathbf{y} \mapsto \langle \mathbf{x}, \mathbf{y} \rangle$ est une fonctionnelle linéaire sur $\mathscr{W}$, i. e.

$$\langle \mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2 \rangle = \langle \mathbf{x}, \mathbf{y}_1 \rangle + \langle \mathbf{x}, \mathbf{y}_2 \rangle,$$
$$\langle \mathbf{x}, k\mathbf{y} \rangle = k \langle \mathbf{x}, \mathbf{y} \rangle$$

pour n'importe quels $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y} \in \mathscr{W}$ et tout $k \in \mathbb{K}$;

c) on associe à tout vecteur $\mathbf{x} \in \mathscr{V}$ un vecteur $\mathbf{y} \in \mathscr{W}$ tel que $\langle \mathbf{x}, \mathbf{y} \rangle \neq 0$, et, inversement, il existe pour tout $\mathbf{y} \in \mathscr{W}$ un vecteur $\mathbf{x} \in \mathscr{V}$ pour lequel $\langle \mathbf{x}, \mathbf{y} \rangle \neq 0$.

Les conditions a) et b) signifient que $\langle \mathbf{x}, \mathbf{y} \rangle$ est *bilinéaire* et la condition c) en exprime la *non-dégénérescence*.

Définition 3. On dit d'une fonction $\mathbf{x}, \mathbf{y} \mapsto \langle \mathbf{x}, \mathbf{y} \rangle$ remplissant les conditions a), b) et c) que c'est un *couplage* des espaces $\mathscr{V}$ et $\mathscr{W}$. Deux espaces $\mathscr{V}$ et $\mathscr{W}$ pour lesquels il existe au moins un couplage sont dits *en dualité* (ce qui est noté $\mathscr{V} \mid \mathscr{W}$).

On observe que la relation de dualité est évidemment symétrique, i.e. *si* $\mathscr{V} \mid \mathscr{W}$, *alors* $\mathscr{W} \mid \mathscr{V}$.

Proposition 3. *Tout espace vectoriel* $\mathscr{V}$ *et son dual* $\mathscr{V}'$ *sont en dualité*:

$$\mathscr{V} \mid \mathscr{V}'.$$

Démonstration. On pose pour tout $\mathbf{x} \in \mathscr{V}$ et tout $\xi \in \mathscr{V}'$:

$$\langle \mathbf{x}, \xi \rangle = \xi(\mathbf{x}).$$

On est évidemment dans les conditions a) et b) (par exemple, $\langle \mathbf{x}, \xi_1 + \xi_2 \rangle = (\xi_1 + \xi_2)(\mathbf{x}) = \xi_1(\mathbf{x}) + \xi_2(\mathbf{x}) = \langle \mathbf{x}, \xi_1 \rangle + \langle \mathbf{x}, \xi_2 \rangle$). L'inégalité $\xi \neq 0$ équivaut à l'existence d'un vecteur $\mathbf{x} \in \mathscr{V}$ tel que $\xi(\mathbf{x}) \neq 0$. Donc $\langle \mathbf{x}, \xi \rangle \neq 0$. De même, l'inégalité

$\mathbf{x} \neq 0$ signifie que $\mathbf{x}^{i_0} \neq 0$ pour au moins un i_0, si bien que $\langle \mathbf{x}, \xi \rangle = \xi(\mathbf{x}) = \mathbf{x}^{i_0} \neq 0$ pour $\xi = \mathbf{e}^{i_0}$. Ainsi, on est dans la condition c). □

La réciproque s'énonce sous forme de

Proposition 4. *Si les espaces $\mathscr{V}$ et $\mathscr{W}$ sont en dualité, chacun d'eux est isomorphe au dual de l'autre:*

$$\mathscr{V} \approx \mathscr{W}', \quad \mathscr{W} \approx \mathscr{V}'.$$

Démonstration. La relation de dualité étant symétrique, il suffit de démontrer le premier isomorphisme. Soit $\mathbf{x} \in \mathscr{V}$. Selon b), la fonction $\mathbf{y} \mapsto \langle \mathbf{x}, \mathbf{y} \rangle$ est une fonctionnelle linéaire sur $\mathscr{W}$, i. e. c'est un vecteur de $\mathscr{W}'$. On désigne cette fonctionnelle par $\varphi(\mathbf{x})$ et on obtient une application

$$\varphi : \mathscr{V} \to \mathscr{W}'.$$

Ainsi, on a par définition:

$$\varphi(\mathbf{x})(\mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle.$$

La condition a) entraîne donc

$$\varphi(\mathbf{x}_1 + \mathbf{x}_2)(\mathbf{y}) = \langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{y} \rangle = \\ = \langle \mathbf{x}_1, \mathbf{y} \rangle + \langle \mathbf{x}_2, \mathbf{y} \rangle = \varphi(\mathbf{x}_1)(\mathbf{y}) + \varphi(\mathbf{x}_2)(\mathbf{y}),$$

i. e.

$$\varphi(\mathbf{x}_1 + \mathbf{x}_2) = \varphi(\mathbf{x}_1) + \varphi(\mathbf{x}_2).$$

De même

$$\varphi(k\mathbf{x})(\mathbf{y}) = \langle k\mathbf{x}, \mathbf{y} \rangle = k \langle \mathbf{x}, \mathbf{y} \rangle = k\varphi(\mathbf{x})(\mathbf{y}),$$

i. e.

$$\varphi(k\mathbf{x}) = k\varphi(\mathbf{x}),$$

ce qui démontre que φ est un homomorphisme.

Si $\varphi(\mathbf{x}) = 0$, alors $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ pour tous les $\mathbf{y} \in \mathscr{W}$, donc $\mathbf{x} = 0$ (en vertu de c)). Ainsi, $\operatorname{Ker} \varphi = \mathbf{0}$. Aussi $\operatorname{Im} \varphi \approx \mathscr{V}$ et, partant, $\dim \mathscr{V} = \dim \operatorname{Im} \varphi \leqslant \dim \mathscr{W}$.

La propriété de symétrie de la relation de dualité fait que chaque fois qu'on a $\dim \mathscr{V} \leqslant \dim \mathscr{W}$, on a nécessairement $\dim \mathscr{W} \leqslant \dim \mathscr{V}$, donc $\dim \mathscr{V} = \dim \mathscr{W}$ et, en particulier, $\dim \operatorname{Im} \varphi = \dim \mathscr{W}$, i. e. $\operatorname{Im} \varphi = \mathscr{W}$. L'homomorphisme φ est donc un isomorphisme. □

* * *

Puisque $\mathscr{V} | \mathscr{V}'$ (proposition 3), on a $\mathscr{V}' | \mathscr{V}$ (propriété de symétrie), donc $\mathscr{V} \approx (\mathscr{V}')'$ (proposition 4). Ce résultat important mérite qu'on l'énonce sous forme de

Théorème 1. *L'espace $(\mathscr{V}')'$ bidual de $\mathscr{V}$ lui est isomorphe:*

$$(\mathscr{V}')' \approx \mathscr{V}. \ \square$$

L'isomorphisme $\mathscr{V} \to (\mathscr{V}')'$ est défini explicitement par la correspondance entre un vecteur $\mathbf{x} \in \mathscr{V}$ et la fonctionnelle $\hat{\mathbf{x}}$ sur $\mathscr{V}'$ donnée par

$$\hat{\mathbf{x}}(\xi) = \xi(\mathbf{x}).$$

La fonctionnelle $\hat{\mathbf{x}}$ est en général identifiée à $\mathbf{x}$, ce qui explique en particulier qu'on la note $\mathbf{x}$ elle aussi.

On dirait que le théorème 1 est une conséquence banale du fait que $\mathscr{V}$ et $(\mathscr{V}')'$ sont de même dimension. Mais il établit en fait un isomorphisme « canonique» entre $\mathscr{V}$ et son bidual, qu'on construit sans aucun arbitraire. D'où la possibilité d'identifier $\hat{\mathbf{x}}$ et $\mathbf{x}$ (donc $(\mathscr{V}')'$ et $\mathscr{V}$).

* * *

Les espaces $\mathscr{V}$ et $\mathscr{V}'$ ont même dimension eux aussi, mais on n'établit en général aucun isomorphisme canonique $\mathscr{V} \to \mathscr{V}'$. Le fait est que certaines notions indispensables nous font défaut pour le moment (nous ne disposons pas par exemple de définition rigoureuse de l'isomorphisme « canonique»), si bien que nous nous limiterons à démontrer l'impossibilité de le construire par le procédé qui paraît le plus simple et le mieux indiqué.

Soient $\mathbf{e}_1, \ldots, \mathbf{e}_n$ une base quelconque de l'espace $\mathscr{V}$ et $\mathbf{e}^1, \ldots, \mathbf{e}^n$ la base duale de $\mathscr{V}'$. Considérons l'isomorphisme $\mathscr{V} \to \mathscr{V}'$ défini par égalité des coordonnées dans ces bases (il associe à chaque vecteur $\mathbf{x} = x^1\mathbf{e}_1 + \ldots + x^n\mathbf{e}_n$ un covecteur $\xi = x^1\mathbf{e}^1 + \ldots + x^n\mathbf{e}^n$ qui est représenté dans la base $\mathbf{e}^1, \ldots, \mathbf{e}^n$ par les mêmes coordonnées que le vecteur $\mathbf{x}$ l'est dans $\mathbf{e}_1, \ldots, \mathbf{e}_n$) et vérifions s'il est « canonique» (i.e. indépendant de la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$). La réponse est négative.

Examinons au préalable le problème général relatif à la transformation des coordonnées des covecteurs dans le passage de la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ à une autre.

Nous utiliserons la convention d'Einstein. Il y a donc intérêt à introduire le *symbole de Kronecker* δ_i^j défini par la formule

$$\delta_i^j = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j. \end{cases}$$

La propriété fondamentale du symbole de Kronecker est fournie par les égalités

$$a^i\delta_i^j = a^j, \quad b_j\delta_i^j = b_i$$

(en effet, les seuls termes non nuls des sommes à gauche sont $a^j \cdot 1 = a^j$ et $b_i \cdot 1 = b_i$ respectivement).

Avec le symbole de Kronecker, on exprime la propriété caractéristique de la base duale par la seule formule

$$\mathbf{e}^j(\mathbf{e}_i) = \delta^j_i.$$

On traduit de même par deux égalités équivalentes la propriété, pour les matrices $(c^{i'}_i)$ et $(c^i_{i'})$, d'être mutuellement réciproques:

$$c^{i'}_i c^j_{i'} = \delta^j_i, \quad c^{i'}_i c^i_{j'} = \delta^{i'}_{j'}.$$

Cela étant, on prend, en plus de $\mathbf{e}_1, \ldots, \mathbf{e}_n$, une base $\mathbf{e}_{1'}, \ldots$ $\ldots, \mathbf{e}_{n'}$ telle que

$$\mathbf{e}_{i'} = c^i_{i'}\mathbf{e}_i, \quad \mathbf{e}_i = c^{i'}_i \mathbf{e}_{i'},$$

avec $C = (c^i_{i'})$ la matrice de passage et $C^{-1} = (c^{i'}_i)$ son inverse. Il est déjà apparu (voir I.11) que les coordonnées x^i et $x^{i'}$ des vecteurs satisfont aux égalités

$$x^i = c^i_{i'} x^{i'}, \quad x^{i'} = c^{i'}_i x^i.$$

Soit $\mathbf{e}^{1'}, \ldots, \mathbf{e}^{n'}$ la duale de $\mathbf{e}_{1'}, \ldots, \mathbf{e}_{n'}$. Par définition,

$$\mathbf{e}^{j'}(\mathbf{e}_{i'}) = \delta^{j'}_{i'}.$$

Donc

$$\mathbf{e}^{j'}(\mathbf{e}_i) = \mathbf{e}^{j'}(c^{i'}_i \mathbf{e}_{i'}) = c^{i'}_i \delta^{j'}_{i'} = c^{j'}_i.$$

Selon la proposition 2,

$$\xi = \xi(\mathbf{e}_i)\,\mathbf{e}^i$$

pour tout covecteur $\xi \in \mathscr{V}'$, si bien qu'on a en particulier

$$\mathbf{e}^{i'} = c^{i'}_i \mathbf{e}^i$$

et, partant,

$$\mathbf{e}^i = c^i_{i'} \mathbf{e}^{i'}$$

(on aboutit à la dernière formule moyennant la propriété de symétrie ou par les calculs: $c^i_{i'}\mathbf{e}^{i'} = c^i_{i'} c^{i'}_j \mathbf{e}^j = \delta^i_j \mathbf{e}^j = \mathbf{e}^i$).

On a de même pour les coordonnées $\xi_i = \xi(\mathbf{e}_i)$ et $\xi_{i'} = \xi(\mathbf{e}_{i'})$ d'un covecteur ξ quelconque:

$$\xi(\mathbf{e}_i) = c^{i'}_i \xi(\mathbf{e}_{i'}),$$

i.e.

$$\xi_i = c^{i'}_i \xi_{i'},$$

et la propriété de symétrie (ou les mêmes calculs) donne

$$\xi_{i'} = c^i_{i'} \xi_i.$$

Ainsi, *les covecteurs de la base duale se transforment comme les coordonnées des vecteurs*, et, corrélativement, *les coordonnées des covecteurs le sont comme les vecteurs de base.*

On dit d'ordinaire que dans un changement de base les vecteurs varient *de façon cogrédiente* et que les coordonnées des vecteurs se transforment *de façon contragrédiente* (c'est la transformation à matrice transposée de l'inverse). Ainsi, *la transformation des bases duales s'effectue de manière contragrédiente et celle des coordonnées des vecteurs de maniére cogrédiente.* □

Il en résulte que si le vecteur $\mathbf{x}$ et le covecteur ξ ont mêmes coordonnées dans une base (et dans sa duale), leurs coordonnées dans une autre base ne coïncident pas (car la transformation des coordonnées des vecteurs et celle des coordonnées des covecteurs sont régies par des formules différentes). Par conséquent, l'application définie par égalité des coordonnées dans les bases duales dépend de la base, i. e. elle n'est pas canonique.

* * *

Soit $S \subset \mathcal{V}$ une partie quelconque de l'espace vectoriel $\mathcal{V}$.

Définition 4. On appelle *annulateur de l'ensemble* S, et on note Ann S ou S°, l'ensemble de toutes les fonctionnelles linéaires $\xi \in \mathcal{V}'$ qui sont nulles sur tout vecteur $\mathbf{x} \in S$.

Ainsi,

$$\operatorname{Ann} S = \{\xi \in \mathcal{V}' ;\ \xi(\mathbf{x}) = 0 \text{ pour tout } \mathbf{x} \in S\}.$$

Evidemment, S° *est un sous-espace de* $\mathcal{V}'$, et $S^\circ \supset T^\circ$ *si* $S \subset T$. □

Proposition 5. *L'annulateur d'un ensemble* $S \subset \mathcal{V}$ *quelconque est identique à l'annulateur de son enveloppe linéaire*:

$$\operatorname{Ann} S = \operatorname{Ann} [S].$$

Démonstration. Comme $S \subset [S]$, on a $S^\circ \supset [S]^\circ$. Inversement, soit $\xi \in S^\circ$. Tout vecteur $k_1\mathbf{x}_1 + \ldots + k_m\mathbf{x}_m$ de $[S]$, où $\mathbf{x}_1, \ldots, \mathbf{x}_m \in S$, vérifie l'égalité

$$\xi(k_1\mathbf{x}_1 + \ldots + k_m\mathbf{x}_m) = k_1\xi(\mathbf{x}_1) + \ldots + k_m\xi(\mathbf{x}_m) = 0$$

par suite de $\xi(\mathbf{x}_1) = 0, \ldots, \xi(\mathbf{x}_m) = 0$. Par conséquent, $\xi \in [S]^\circ$, i.e. $S^\circ \subset [S]^\circ$. □

Cette proposition autorise à se borner au cas de sous-espaces.

Il est clair que $\operatorname{Ann} \mathbf{0} = \mathcal{V}'$. Inversement, *si* $\operatorname{Ann} S = \mathcal{V}'$, *alors* $S = \{\mathbf{0}\}$ (car $\mathbf{x} = \mathbf{0}$ si $\xi(\mathbf{x}) = 0$ pour tous les $\xi \in \mathcal{V}'$). □

De même, $\operatorname{Ann} \mathcal{V} = \mathbf{0}$. Inversement, *si* $\operatorname{Ann} S = \mathbf{0}$, *alors* $[S] = \mathcal{V}$. En effet, si $[S] \neq \mathcal{V}$ et si $\mathbf{e}_1, \ldots, \mathbf{e}_n$ est une base de $\mathcal{V}$ telle que $[S] = [\mathbf{e}_1, \ldots, \mathbf{e}_m]$, $m < n$, alors $\mathbf{e}^n \in [S]^\circ$, donc $S^\circ \neq \mathbf{0}$. □

Proposition 6. *Etant donné un sous-espace* $\mathcal{P} \subset \mathcal{V}$ *quelconque, on a l'égalité*

$$\dim \mathcal{P}^\circ = n - \dim \mathcal{P}.$$

Démonstration. Soient $\dim \mathcal{P} = p$ et $\mathbf{e}_1, \ldots, \mathbf{e}_p, \ldots, \mathbf{e}_n$ une base de l'espace $\mathcal{V}$ telle que $\mathcal{P} = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$. Considérons la base duale

$$\mathbf{e}^1, \ldots, \mathbf{e}^p, \ldots, \mathbf{e}^n.$$

Si $i \leqslant p$ et $j > p$, on a manifestement $i \neq j$, donc $\mathbf{e}^j(\mathbf{e}_i) = 0$. Par conséquent, $\mathbf{e}^{p+1}, \ldots, \mathbf{e}^n \in [\mathbf{e}_1, \ldots, \mathbf{e}_p]^\circ = \mathcal{P}^\circ$. D'autre part, $\xi \in \mathcal{P}^\circ$ entraîne $\xi(\mathbf{e}_1) = 0, \ldots, \xi(\mathbf{e}_p) = 0$, donc $\xi = \xi_{p+1}\mathbf{e}^{p+1} + \ldots + \xi_n\mathbf{e}^n$.

Ainsi, *les covecteurs* $\mathbf{e}^{p+1}, \ldots, \mathbf{e}^n$ *constituent une base du sous-espace* $\mathcal{P}^\circ$. Par conséquent, $\dim \mathcal{P}^\circ = n - p$. □

Puisque $\mathcal{V} = (\mathcal{V}')'$ par le théorème 1, on remplace partout $\mathcal{V}$ par $\mathcal{V}'$, et vice versa. En particulier, on définit pour tout ensemble $S \subset \mathcal{V}'$ un sous-espace $\operatorname{Ann} S \subset \mathcal{V}$ formé de vecteurs $\mathbf{x} \in \mathcal{V}$ tels que $\mathbf{x}(\xi) = 0$ (i.e. $\xi(\mathbf{x}) = 0$) pour tout covecteur $\xi \in S$, et la dimension du sous-espace est $n - r$, r étant la dimension de $[S]$, i. e. le rang de l'ensemble S.

Ainsi, on a montré que, premièrement, on définit les sous-espaces de $\mathcal{V}$ non seulement comme des enveloppes linéaires, mais aussi au sens « dual » en tant qu'annulateurs de l'ensemble de covecteurs $S = \{\xi_1, \ldots, \xi_m\}$, i.e. par des équations de la forme

$$\xi_1(\mathbf{x}) = 0, \ldots, \xi_m(\mathbf{x}) = 0, \tag{5}$$

et, deuxièmement, on peut calculer efficacement la dimension du sous-espace ainsi défini (cette dimension est égale à $n - r$, avec r le rang de $\{\xi_1, \ldots, \xi_m\}$).

* * *

Il y a intérêt à énoncer ces résultats en termes de coordonnées.

En coordonnées, les covecteurs $\xi_1, \ldots, \xi_m$ sont exprimés par les formes linéaires des $x^1, \ldots, x^n$ (proposition 1), si bien que les équations (5) se récrivent

$$\begin{aligned} a_{11}x^1 + \ldots + a_{1n}x^n &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x^1 + \ldots + a_{mn}x^n &= 0, \end{aligned} \tag{6}$$

i.e. ce sont des équations linéaires homogènes usuelles. On a donc le

Théorème 2. *L'ensemble des solutions* $(x^1, \ldots, x^n)$ *du système* (6) *d'équations linéaires homogènes est un sous-espace de dimension* $n - r$ *de l'espace* $\mathbb{K}^n$, r *étant le rang de la matrice des coefficients*

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}. \ \square$$

On cherche une base de ce sous-espace, i. e. $n - r$ solutions linéairement indépendantes (qui constituent par définition le *système*

fondamental de solutions) par le procédé décrit dans la leçon 2, qui consiste à résoudre (6) en faisant prendre à $n - r$ inconnues « libres » $n - r$ jeux de valeurs en veillant à ce que les solutions obtenues soient linéairement indépendantes. Pour cela il suffit de choisir les jeux de valeurs de façon qu'ils forment une matrice carrée d'ordre $n - r$ qui soit non dégénérée (il y a avantage à les choisir tels qu'ils constituent une matrice unité).

LEÇON 5

Annulateur de l'annulateur et annulateurs de termes directs. — Fonctionnelles bilinéaires et formes bilinéaires. — Fonctionnelles bilinéaires dans l'espace dual. — Fonctionnelles bilinéaires mixtes. — Tenseurs.

Les annulateurs étant définis pour les parties de l'espace $\mathscr{V}'$, on peut parler de l'annulateur de l'annulateur

$$\text{Ann Ann } S = S^{\circ\circ}$$

d'une partie S quelconque de $\mathscr{V}$.

Proposition 1. *Pour tout sous-espace $\mathscr{P}$ de $\mathscr{V}$ on a l'égalité*

$$\mathscr{P}^{\circ\circ} = \mathscr{P}.$$

Démonstration. Si $\mathbf{x} \in \mathscr{P}$, alors $\xi(\mathbf{x}) = 0$ pour tout $\xi \in \mathscr{P}^\circ$, i. e. $\mathbf{x}(\xi) = 0$. Cela veut dire que $\mathbf{x} \in \mathscr{P}^{\circ\circ}$. Ainsi, $\mathscr{P}^{\circ\circ} \subset \mathscr{P}$, donc $\mathscr{P}^{\circ\circ} = \mathscr{P}$ car $\dim \mathscr{P}^{\circ\circ} = n - \dim \mathscr{P}^\circ = n - (n - \dim \mathscr{P}) = \dim \mathscr{P}$. □

Lorsque S est un ensemble arbitraire, on a évidemment $S^{\circ\circ} = [S]$.

Proposition 2. *Si $\mathscr{V} = \mathscr{P} \oplus \mathscr{Q}$, alors $\mathscr{V}' = \mathscr{P}^\circ \oplus \mathscr{Q}^\circ$, et $\mathscr{P}^\circ \approx \mathscr{Q}'$, $\mathscr{Q}^\circ \approx \mathscr{P}'$.*

Démonstration. Soit $\dim \mathscr{P} = p$ et $\dim \mathscr{Q} = q$, auquel cas $p + q = n$ et $\mathscr{P} \cap \mathscr{Q} = \mathbf{0}$. Aussi $\dim \mathscr{P}^\circ + \dim \mathscr{Q}^\circ = (n - p) + (n - q) = n$. De plus, si $\xi \in \mathscr{P}^\circ \cap \mathscr{Q}^\circ$, alors $\xi(\mathbf{x}) = 0$ pour tout $\mathbf{x} \in \mathscr{P}$ et $\xi(\mathbf{y}) = 0$ pour tout $\mathbf{y} \in \mathscr{Q}$. Aussi $\xi(\mathbf{x} + \mathbf{y}) = 0$, donc $\xi(z) = 0$ pour tout $z \in \mathscr{V}$. Par conséquent, $\xi = \mathbf{0}$, i.e. $\mathscr{P}^\circ \cap \mathscr{Q}^\circ = \mathbf{0}$, ce qui démontre $\mathscr{V}' = \mathscr{P}' \oplus \mathscr{Q}^\circ$ (voir leçon 3, conséquence de la prop. 3).

On associe à chaque fonctionnelle linéaire $\xi \in \mathscr{P}^\circ$ sa restriction

$$\xi' = \xi|_{\mathscr{Q}}$$

au sous-espace $\mathscr{Q}$. On obtient une application $\xi \to \xi'$ de l'espace $\mathscr{P}^\circ$ dans l'espace $\mathscr{Q}'$ qui est visiblement linéaire (c'est un homomorphisme). Le noyau de l'application est formé de toutes les fonctionnelles $\xi \in \mathscr{P}^\circ$ telles que $\xi|_{\mathscr{Q}} = 0$, i.e. telles que $\xi \in \mathscr{Q}^\circ$. Or, on sait que $\mathscr{P}^\circ \cap \mathscr{Q}^\circ = \mathbf{0}$, si bien que $\xi \mapsto \xi'$ est un monomorphisme.

Soit $\eta \in \mathscr{Q}'$. On définit sur $\mathscr{V}$ une fonctionnelle ξ en posant

$$\xi(\mathbf{x} + \mathbf{y}) = \eta(\mathbf{y})$$

pour tout vecteur $\mathbf{x} + \mathbf{y}$, $\mathbf{x} \in \mathcal{P}$, $\mathbf{y} \in \mathcal{Q}$. Il est clair que ξ est bien définie, linéaire, appartient à $\mathcal{P}^\circ$ et que $\xi' = \eta$, ce qui prouve que $\xi \mapsto \xi'$ est un isomorphisme.

L'isomorphisme $\mathcal{Q}^\circ \approx \mathcal{P}'$ est démontré de façon analogue. □

On note que les isomorphismes de la proposition 2 sont « canoniques ».

* * *

Définition 1. Une fonction $B: \mathbf{x}, \mathbf{y} \mapsto B(\mathbf{x}, \mathbf{y}) \in \mathbb{K}$ de *deux* variables vectorielles $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ s'appelle *fonctionnelle bilinéaire sur* $\mathcal{V}$ si, lorsqu'on fixe chacune des variables, elle est linéaire par rapport à l'autre, i.e.

$$B(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = B(\mathbf{x}_1, \mathbf{y}) + B(\mathbf{x}_2, \mathbf{y}),$$
$$B(k\mathbf{x}, \mathbf{y}) = kB(\mathbf{x}, \mathbf{y})$$

et

$$B(\mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2) = B(\mathbf{x}, \mathbf{y}_1) + B(\mathbf{x}, \mathbf{y}_2),$$
$$B(\mathbf{x}, k\mathbf{y}) = kB(\mathbf{x}, \mathbf{y})$$

quels que soient les vecteurs $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}, \mathbf{y}_1, \mathbf{y}_2, \mathbf{y} \in \mathcal{V}$ et pour tout $k \in \mathbb{K}$.

Un exemple en est le produit scalaire $(\mathbf{x}, \mathbf{y})$ (voir I.13). Il en est de même des couplages introduits dans la leçon 4 encore que leurs variables appartiennent en général à deux espaces différents. La théorie ci-dessous se généralise en principe à ce cas, mais les raisonnements nécessaires sont suffisamment fastidieux pour qu'on passe outre.

Soit $\mathbf{e}_1, \ldots, \mathbf{e}_n$ une base arbitraire de l'espace $\mathcal{V}$. On pose

$$b_{ij} = B(\mathbf{e}_i, \mathbf{e}_j), \tag{1}$$

et on a pour $\mathbf{x} = x^i\mathbf{e}_i$ et $\mathbf{y} = y^j\mathbf{e}_j$ quelconques

$$B(\mathbf{x}, \mathbf{y}) = B(\mathbf{e}_i, \mathbf{e}_j)\, x^i y^j = b_{ij} x^i y^j.$$

Cela démontre la chaîne d'égalités

$$B(\mathbf{x}, \mathbf{y}) = b_{ij}x^iy^j =$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x^iy^j = b_{11}x^1y^1 + \ldots + b_{1n}x^1y^n + \tag{2}$$
$$+ b_{21}x^2y^1 + \ldots + b_{2n}x^2y^n + b_{n1}x^ny^1 + \ldots + b_{nn}x^ny^n.$$

L'expression algébrique du second membre s'appelle (c'est connu dès I.14) la *forme bilinéaire* en $x^1, \ldots, x^n$ et $y^1, \ldots, y^n$. Ainsi, *l'expression en coordonnées de toute fonctionnelle bilinéaire est une forme bilinéaire* à coefficients (1) (qu'on appelle, pour simplifier, les *coefficients de la fonctionnelle B*). Inversement, toute forme bilinéaire

définit par la formule (2) (on le constate facilement) une fonctionnelle bilinéaire. Il y a donc (pour la base donnée !) une bijection entre les fonctionnelles bilinéaires et les formes bilinéaires.

Les coefficients (1) de B constituent la matrice

$$B=\begin{pmatrix} b_{11} & \dots & b_{1n} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{nn} \end{pmatrix}$$

dite *matrice de la fonctionnelle bilinéaire* B (par rapport à la base donnée).

On conçoit que la somme de deux fonctionnelles bilinéaires est une fonctionnelle bilinéaire, ainsi que le produit d'une fonctionnelle bilinéaire par un nombre. On dit donc que *l'ensemble* $\mathsf{T}_2(\mathscr{V})$ *des fonctionnelles bilinéaires sur* $\mathscr{V}$ *est un espace vectoriel.* □

Lorsqu'on fait la somme de deux fonctionnelles bilinéaires, leurs matrices s'ajoutent, et si l'on multiplie une fonctionnelle bilinéaire par un nombre, sa matrice l'est également. Cela signifie que *la correspondance qui associe à une fonctionnelle bilinéaire sa matrice est un isomorphisme de l'espace vectoriel* $\mathsf{T}_2(\mathscr{V})$ *sur l'espace vectoriel des matrices carrées d'ordre* n. □

La matrice B aidant, on récrit (2):

$$B(\mathbf{x}, \mathbf{y}) = x^{\mathsf{T}}By,$$

avec, une fois de plus,

$$x=\begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}, \quad y=\begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix}.$$

(Cf. I.14, où l'on a obtenu des formules analogues du produit scalaire.)

Soient $\xi, \eta \in \mathsf{T}_1(\mathscr{V})$ deux fonctionnelles linéaires. La formule

$$(\xi \otimes \eta)(\mathbf{x}, \mathbf{y}) = \xi(\mathbf{x})\,\eta(\mathbf{y})$$

définit évidemment une fonctionnelle bilinéaire $\xi \otimes \eta$.

Définition 2. La fonctionnelle $\xi \otimes \eta$ s'appelle le *produit tensoriel* des fonctionnelles ξ et η.

Considérons en particulier les produits tensoriels $\mathbf{e}^i \otimes \mathbf{e}^j$ des covecteurs de la base duale. Comme $\mathbf{e}^i(\mathbf{x}) = x^i$ et $\mathbf{e}^i(\mathbf{y}) = y^j$, on a

$$(\mathbf{e}^i \otimes \mathbf{e}^j)(\mathbf{x}, \mathbf{y}) = x^i y^j,$$

et la fonctionnelle $B' = b_{ij}(\mathbf{e}^i \otimes \mathbf{e}^j)$ vérifie la formule

$$B'(\mathbf{x}, \mathbf{y}) = b_{ij}(\mathbf{e}^i \otimes \mathbf{e}^j)(\mathbf{x}, \mathbf{y}) = b_{ij}x^i y^j. \tag{3}$$

En particulier, $B'(\mathbf{e}_i, \mathbf{e}_j) = b_{ij}$, d'où l'indépendance linéaire des fonctionnelles bilinéaires $\mathbf{e}^i \otimes \mathbf{e}^j$, $i, j = 1, \dots, n$ (si $B' = 0$,

alors $b_{ij} = 0$). Soit maintenant $B \in \mathsf{T}_2(\mathscr{V})$ quelconque. Si l'on forme B' à partir des coefficients b_{ij} de B, on a $B' = B$ selon (3).

On vient de démontrer la

Proposition 3. *Les produits tensoriels*

$$\mathbf{e}^i \otimes \mathbf{e}^j, \quad i, j = 1, \ldots, n,$$

des vecteurs de la base duale forment une base de l'espace vectoriel $\mathsf{T}_2(\mathscr{V})$. *Les coordonnées d'une fonctionnelle bilinéaire* $B \in \mathsf{T}_2(\mathscr{V})$ *quelconque dans cette base sont ses coefficients* b_{ij}:

$$B = b_{ij}\mathbf{e}^i \otimes \mathbf{e}^j. \ \square$$

En particulier,

$$\dim \mathsf{T}_2(\mathscr{V}) = n^2.$$

Passons à une nouvelle base:

$$\mathbf{e}_{i'} = c^i_{i'}\mathbf{e}_i.$$

On a

$$b_{i'j'} = B(\mathbf{e}_{i'}, \mathbf{e}_{j'}) = c^i_{i'}c^j_{j'}B(\mathbf{e}_i, \mathbf{e}_j) = c^i_{i'}c^j_{j'}b_{ij}.$$

Ainsi, les coefficients dans la nouvelle base de la fonctionnelle bilinéaire B sont donnés par la formule

$$b_{i'j'} = c^i_{i'}c^j_{j'}b_{ij}$$

qui s'écrit en notations matricielles

$$B' = C^{\mathsf{T}}BC,$$

où $B = (b_{ij})$, $B' = (b_{i'j'})$ et $C = (c^i_{i'})$ (cf. I.14).

* * *

Les fonctionnelles bilinéaires $B: \xi, \eta \mapsto B(\xi, \eta)$ *des covecteurs* $\xi, \eta \in \mathscr{V}'$ se traitent de façon analogue. La seule différence est la position des indices. Les valeurs de la fonctionnelle sont données par la formule

$$B(\xi, \eta) = b^{ij}\xi_i\eta_j,$$

où $b^{ij} = B(\mathbf{e}^i, \mathbf{e}^j)$ et $\xi_i = \xi(\mathbf{e}_i)$ et $\eta_j = \eta(\mathbf{e}_j)$ sont les coordonnées des covecteurs ξ et η. Si l'on passe à la base $\mathbf{e}_{i'} = c^i_{i'}\mathbf{e}_i$, on a

$$b^{i'j'} = c^{i'}_i c^{j'}_j b^{ij}.$$

Les fonctionnelles bilinéaires des covecteurs constituent l'espace vectoriel

$$\mathsf{T}^2(\mathscr{V}) = \mathsf{T}_2(\mathscr{V}')$$

de dimension n^2. *Le produit tensoriel* $\mathbf{x} \otimes \mathbf{y}$ des vecteurs $\mathbf{x}$ et $\mathbf{y}$ est par définition une fonctionnelle de $\mathsf{T}^2(\mathscr{V})$ définie par l'égalité

$$(\mathbf{x} \otimes \mathbf{y})(\xi, \eta) = \xi(\mathbf{x})\,\eta(\mathbf{y}).$$

Les produits tensoriels

$$\mathbf{e}_i \otimes \mathbf{e}_j, \qquad i, j = 1, \ldots, n,$$

forment une base de l'espace $\mathsf{T}^2(\mathscr{V})$, et

$$B = b^{ij}\, \mathbf{e}_i \otimes \mathbf{e}_j$$

pour toute B de $\mathsf{T}^2(\mathscr{V})$.

* * *

Le cas des fonctionnelles bilinéaires

$$B : \mathbf{x}, \xi \mapsto B(\mathbf{x}, \xi),$$

$\mathbf{x}$ étant un vecteur de $\mathscr{V}$ et ξ un covecteur de $\mathscr{V}'$, est plus intéressant. Nous dirons de ces fonctionnelles qu'elles sont *mixtes*. Les fonctionnelles mixtes constituent également un espace vectoriel de dimension n^2 que nous désignerons par $\mathsf{T}^1_1(\mathscr{V})$.

En coordonnées, les valeurs d'une fonctionnelle mixte B s'écrivent

$$B(\mathbf{x}, \xi) = b^j_i x^i \xi_j,$$

où $b^j_i = B(\mathbf{e}_i, \mathbf{e}^j)$ et $x^i = \mathbf{e}^i(\mathbf{x})$, $\xi_j = \xi(\mathbf{e}_j)$ sont les coordonnées du vecteur $\mathbf{x}$ et du covecteur ξ (dans les bases duales $\mathbf{e}_1, \ldots, \mathbf{e}_n$ et $\mathbf{e}^1, \ldots, \mathbf{e}^n$).

On définit le produit tensoriel $\eta \otimes \mathbf{y}$ d'un covecteur η et d'un vecteur $\mathbf{y}$ par

$$(\eta \otimes \mathbf{y})(\mathbf{x}, \xi) = \eta(\mathbf{x})\, \xi(\mathbf{y}),$$

et on a de suite que *les produits tensoriels* $\mathbf{e}^i \otimes \mathbf{e}_j$ *forment une base de l'espace* $\mathsf{T}^1_1(\mathscr{V})$, *et*

$$B = b^j_i \mathbf{e}^i \otimes \mathbf{e}_j$$

pour toute $B \in \mathsf{T}^1_1(\mathscr{V})$. □

Les coefficients $b^{j'}_{i'}$ de $B \in \mathsf{T}^1_1(\mathscr{V})$ s'expriment dans la base

$$\mathbf{e}_{i'} = c^i_{i'} \mathbf{e}_i$$

par la formule

$$b^{j'}_{i'} = c^i_{i'} c^{j'}_j b^j_i. \tag{4}$$

Il s'agit d'une tout autre transformation que celle qu'on a fait subir aux coefficients des fonctionnelles bilinéaires de $\mathsf{T}_2(\mathscr{V})$ ou de $\mathsf{T}^2(\mathscr{V})$. Pour en avoir une idée intuitive, on va la déduire en notations matricielles.

Soit

$$B = \begin{pmatrix} b^1_1 & \ldots & b^1_n \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ b^n_1 & \ldots & b^n_n \end{pmatrix}$$

et

$$x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}, \quad \xi = (\xi_1, \ldots, \xi_n).$$

Dans ce cas, on constate sans peine que

$$B(\mathbf{x}, \xi) = \xi Bx.$$

Soit, corrélativement,

$$B' = \begin{pmatrix} b_{1'}^{1'} & \ldots & b_{n'}^{1'} \\ \cdot & \cdot & \cdot \\ b_{1}^{n'} & \ldots & b_{n'}^{n'} \end{pmatrix}$$

et

$$x' = \begin{pmatrix} x^{1'} \\ \vdots \\ x^{n'} \end{pmatrix}, \quad \xi' = (\xi_{1'}, \ldots, \xi_{n'}),$$

d'où

$$B(\mathbf{x}, \xi) = \xi' B' x'.$$

On sait que

$$x = Cx'$$

et

$$\xi' = \xi C, \quad \text{i. e.} \quad \xi = \xi' C^{-1}$$

(les coordonnées des covecteurs varient de manière cogrédiente). Aussi

$$\xi Bx = \xi' C^{-1} BCx' = \xi' B' x',$$

i. e.

$$B' = C^{-1} BC.$$

C'est bien la formule (4) en notations matricielles. La matrice inverse C^{-1} remplace ici la transposée $C^{\top}$.

Etant donnés ces résultats, une généralisation vient d'elle-même.

* * *

Définition 3. On appelle *tenseur de type* (p, q), $p, q \geqslant 0$, sur l'espace vectoriel $\mathscr{V}$ une fonction arbitraire

$$T: \mathbf{x}_1, \ldots, \mathbf{x}_p, \xi^1, \ldots, \xi^q \mapsto T(\mathbf{x}_1, \ldots, \mathbf{x}_p, \xi^1, \ldots, \xi^q)$$

de p vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_p$ et q covecteurs $\xi^1, \ldots, \xi^q$ qui est linéaire en chacune des variables (pour les valeurs fixes des autres), i.e. une fonction *multilinéaire*.

On peut donc dire que les fonctionnelles bilinéaires des vecteurs sont des tenseurs de type (2, 0), les fonctionnelles bilinéaires des covecteurs des tenseurs de type (0, 2) et les fonctionnelles bilinéaires mixtes des tenseurs de type (1, 1).

De même, les covecteurs sont des tenseurs de type (1, 0) et les vecteurs des tenseurs de type (0, 1) (par suite de $\mathcal{V} = (\mathcal{V}')'$).

Avec les conventions générales sur les fonctions, on identifie les tenseurs de type (0, 0) sans variables et les éléments du corps $\mathbb{K}$.

L'ensemble des tenseurs de type (p, q) est noté $\mathsf{T}_p^q(\mathcal{V})$, et on omet l'indice $= 0$. Cette notation concorde avec les appellations $\mathsf{T}_2(\mathcal{V})$, $\mathsf{T}^2(\mathcal{V})$ et $\mathsf{T}_1^1(\mathcal{V})$ données aux espaces de fonctionnelles bilinéaires et avec le symbole $\mathsf{T}_1(\mathcal{V})$ qui désigne le dual $\mathcal{V}'$. Conformément aux résultats ci-dessus, $\mathsf{T}^1(\mathcal{V}) = \mathcal{V}$ et $\mathsf{T}_0^0(\mathcal{V}) = \mathbb{K}$.

Il est clair que *tous les ensembles* $\mathsf{T}_p^q(\mathcal{V})$ *constituent des espaces vectoriels* (pour les opérations linéaires ordinaires sur les fonctions).

Soient $\mathbf{e}_1, \ldots, \mathbf{e}_n$ une base quelconque de $\mathcal{V}$ et $\mathbf{e}^1, \ldots, \mathbf{e}^n$ la base duale de $\mathcal{V}'$. Etant donnés

$$\mathbf{x}_1 = x_1^{i_1}\mathbf{e}_{i_1}, \ \ldots, \ \mathbf{x}_p = x_p^{i_p}\mathbf{e}_{i_p},$$
$$\xi^1 = \xi_{j_1}^1\mathbf{e}^{j_1}, \ \ldots, \ \xi^q = \xi_{j_q}^q\mathbf{e}^{j_q}$$

et la propriété « un tenseur de type (p, q) est une fonction multilinéaire », on a

$$T(\mathbf{x}_1, \ldots, \mathbf{x}_p, \xi^1, \ldots, \xi^q) = T_{i_1 \ldots i_p}^{j_1 \ldots j_q} x_1^{i_1} \ldots x_p^{i_p} \xi_{j_1}^1 \ldots \xi_{j_q}^q, \tag{5}$$

où

$$T_{i_1 \ldots i_p}^{j_1 \ldots j_q} = T(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_p}, \mathbf{e}^{j_1}, \ldots, \mathbf{e}^{j_q}). \tag{6}$$

Les nombres $T_{i_1 \ldots i_p}^{j_1 \ldots j_q}$ s'appellent les *coefficients* du tenseur T. Ils sont au nombre de n^{p+q}.

On simplifie les formules par recours aux *multi-indices*

$$\alpha = (i_1, \ldots, i_p) \text{ et } \beta = (j_1, \ldots, j_q).$$

On pose

$$T_\alpha^\beta = T_{i_1, \ldots, i_p}^{j_1, \ldots, j_q}$$

et

$$x^\alpha = x_1^{i_1} \ldots x_p^{i_p}, \quad \xi_\beta = \xi_{j_1}^1 \ldots \xi_{j_q}^q,$$

ce qui permet de condenser (5) :

$$T(\mathbf{x}_1, \ldots, \mathbf{x}_p, \xi^1, \ldots, \xi^q) = T_\alpha^\beta x^\alpha \xi_\beta. \tag{7}$$

La formule (7) signifie qu'en coordonnées tout tenseur est représenté par une *forme multilinéaire* $T_\alpha^\beta x^\alpha \xi_\beta$.

Inversement, toute forme multilinéaire $T^{\beta}_{\alpha}x^{\alpha}\xi_{\beta} = T^{j_1\ldots j_q}_{i_1\ldots i_p}x^{i_1}_1 \ldots$ $\ldots x^{i_p}_p\xi^1_{j_1} \ldots \xi^q_{j_q}$ définit par (7) une fonctionnelle

$$T: \mathbf{x}_1, \ldots, \mathbf{x}_p, \xi^1, \ldots, \xi^q \mapsto T(\mathbf{x}_1, \ldots, \mathbf{x}_p, \xi^1, \ldots, \xi^q)$$

qui est évidemment multilinéaire, i. e. c'est un tenseur.

Ainsi, *étant donnée une base* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *de l'espace* $\mathscr{V}$, *il existe une bijection entre les tenseurs de type* (p, q) *et les formes multilinéaires* $T^{\beta}_{\alpha}x^{\alpha}\xi_{\beta}$, i. e. les ensembles $(T^{\beta}_{\alpha}) = (T^{j_1\ldots j_q}_{i_1\ldots i_p})$ d'éléments de K. □

On passe de $\mathbf{e}_1, \ldots, \mathbf{e}_n$ à une nouvelle base $\mathbf{e}_{1'}, \ldots, \mathbf{e}_{n'}$. Soit

$$\mathbf{e}_{i'} = c^{i}_{i'}\mathbf{e}_i.$$

Le fait que tout tenseur de type (p, q) est linéaire en chacune des variables entraîne pour les coefficients

$$T^{j'_1\ldots j'_q}_{i'_1\ldots i'_p} = T(\mathbf{e}_{i'_1}, \ldots, \mathbf{e}_{i'_p}, \mathbf{e}^{j'_1}, \ldots, \mathbf{e}^{j'_q})$$

de T dans la base $\mathbf{e}_{1'}, \ldots, \mathbf{e}_{n'}$:

$$T^{j'_1\ldots j'_q}_{i'_1\ldots i'_p} = c^{i_1}_{i'_1}\ldots c^{i_p}_{i'_p}c^{j'_1}_{j_1}\ldots c^{j'_q}_{j_q}T^{j_1\ldots j_q}_{i_1\ldots i_p}. \tag{8}$$

C'est la *loi de transformation* des tenseurs. Par convention, nous dirons que chaque indice de la formule (8) varie indépendamment des autres, les indices inférieurs variant de manière cogrédiente et les indices supérieurs de manière contragrédiente.

En notations simplifiées, la formule (8) s'écrit

$$T^{\beta'}_{\alpha'} = c^{\alpha}_{\alpha'}c^{\beta'}_{\beta}T^{\beta}_{\alpha},$$

où

$$c^{\alpha}_{\alpha'} = c^{i_1}_{i'_1}\ldots c^{i_p}_{i'_p}, \quad c^{\beta'}_{\beta} = c^{j'_1}_{j_1}\ldots c^{j'_q}_{j_q}.$$

Théorème 1. *On suppose qu'on associe à chaque base* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *de l'espace* $\mathscr{V}$ n^{p+q} *nombres* $T^{j_1\ldots j_q}_{i_1\ldots i_p}$, *les nombres associés à des bases différentes étant reliés par la loi de transformation* (8). *Alors il existe sur* $\mathscr{V}$ *un seul tenseur de type* (p, q) *dont les coefficients dans chaque base* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *sont* $T^{j_1\ldots j_q}_{i_1\ldots i_p}$ *donnés.*

Démonstration. Il est déjà connu que dès qu'on fixe $T^{j_1\ldots j_q}_{i_1\ldots i_p} = T^{\beta}_{\alpha}$ dans une base donnée $\mathbf{e}_1, \ldots, \mathbf{e}_n$, on détermine moyennant la formule

$$T(\mathbf{x}_1, \ldots, \mathbf{x}_p, \xi^1, \ldots, \xi^q) = T^{j_1\ldots j_q}_{i_1\ldots i_p}x^{i_1}_1 \ldots x^{i_p}_p\xi^1_{j_1}\ldots \xi^q_{j_q} = T^{\beta}_{\alpha}x^{\alpha}\xi_{\beta}$$

un tenseur T de type (p, q). Il suffit donc de vérifier que les coefficients de T ne changent pas avec le passage à toute autre base $\mathbf{e}_{1'}, \ldots$ $\ldots, \mathbf{e}_{n'}$. La chose est évidente. En effet, il est déjà apparu que les coefficients de T dans $\mathbf{e}_{1'}, \ldots, \mathbf{e}_{n'}$ sont les nombres $c^{\alpha}_{\alpha'} c^{\beta'}_{\beta} T^{\beta}_{\alpha}$ qui sont égaux à $T^{\beta'}_{\alpha'}$ par hypothèse. □

D'après ce théorème, on identifie les tenseurs et les ensembles de nombres $(T^{j_1 \ldots j_q}_{i_1 \ldots i_p})$ reliés par les formules (6). C'est justement ces tenseurs qu'on rencontre en général en physique. Dans ce cas, les nombres $T^{j_1 \ldots j_q}_{i_1 \ldots i_p}$ s'appellent les *composantes* et non les coefficients.

LEÇON 6

Produit de tenseurs. — Base de l'espace des tenseurs. — Contraction. — Espace associé à une fonctionnelle multilinéaire.

Nous avons dit dans la leçon 5 qu'on peut additionner les tenseurs de même type. Il est clair que dans cette opération les coefficients (les composantes) des tenseurs s'ajoutent :

$$(T+S)_{i_1\ldots i_p}^{j_1\ldots j_q}=T_{i_1\ldots i_p}^{j_1\ldots j_q}+S_{i_1\ldots i_p}^{j_1\ldots j_q}.$$

Si deux tenseurs s'interprètent comme ensembles de nombres $T_{i_1\ldots i_p}^{j_1\ldots j_q}$, cette formule définit leur somme.

En plus de l'addition, on définit pour les tenseurs l'*opération de multiplication* notée $\otimes$. Le produit d'un tenseur de type (p, q) et d'un autre de type (r, s) est un tenseur de type $(p+r, q+s)$. En passant aux composantes, le produit est donné par

$$(T\otimes S)_{i_1\ldots i_{p+r}}^{j_1\ldots j_{q+s}}=T_{i_1\ldots i_p}^{j_1\ldots j_q}S_{i_{p+1}\ldots i_{p+r}}^{j_{q+1}\ldots j_{q+s}}$$

(ainsi, chaque composante de T est multipliée par chaque composante de S). S'agissant de T et S considérés comme fonctions multilinéaires, on a la formule

$$(T\otimes S)(\mathrm{x}_1, \ldots, \mathrm{x}_{p+r}, \xi^1, \ldots, \xi^{q+s})=$$
$$=T(\mathrm{x}_1, \ldots, \mathrm{x}_p, \xi^1, \ldots, \xi^q)\,S(\mathrm{x}_{p+1}, \ldots, \mathrm{x}_{p+r}, \xi^{q+1}, \ldots, \xi^{q+s}).$$

La multiplication $\otimes$ est évidemment distributive par rapport à l'addition :

$$(T+S)\otimes R=T\otimes R+S\otimes R,$$
$$R\otimes(T+S)=R\otimes T+R\otimes S,$$

et associative :

$$(T\otimes S)\otimes R=T\otimes(S\otimes R).$$

Mais elle est en général non commutative :

$$T\otimes S\neq S\otimes T.$$

Si l'un des facteurs (ou les deux) est un tenseur de type (0, 0), i. e. un élément de K, alors le produit ne diffère pas du produit usuel :

$$k \otimes T = T \otimes k = kT.$$

Tous les espaces vectoriels $\mathsf{T}_p^q(\mathcal{V})$ sont, pour les opérations $+$ et $\otimes$, un objet algébrique qui constitue un exemple d'*algèbre bigraduée*. Cette algèbre est désignée par le symbole $\mathsf{T}(\mathcal{V})$ et porte le nom d'*algèbre tensorielle* de l'espace vectoriel $\mathcal{V}$.

* * *

Soit, une fois de plus, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ une base quelconque de $\mathcal{V}$ et $\mathbf{e}^1, \ldots, \mathbf{e}^n$ la base duale de $\mathcal{V}'$.

On pose pour tout multi-indice $\alpha = (i_1, \ldots, i_p)$ et tout multi-indice $\beta = (j_1, \ldots, j_q)$:

$$\mathbf{e}^\alpha = \mathbf{e}^{i_1} \otimes \ldots \otimes \mathbf{e}^{i_p}, \quad \mathbf{e}_\beta = \mathbf{e}_{j_1} \otimes \ldots \otimes \mathbf{e}_{j_q},$$

auquel cas

$$\mathbf{e}^\alpha(\mathbf{x}_1, \ldots, \mathbf{x}_p) = x_1^{i_1} \ldots x_p^{i_p} = x^\alpha$$

et

$$\mathbf{e}_\beta(\xi^1, \ldots, \xi^q) = \xi_{j_1}^1 \ldots \xi_{j_q}^q = \xi_\beta$$

pour n'importe quels vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_p$ et covecteurs $\xi^1, \ldots, \xi^q$. Aussi les nombres T_α^β quelconques vérifient l'égalité

$$(T_\alpha^\beta \mathbf{e}^\alpha \otimes \mathbf{e}_\beta)(\mathbf{x}_1, \ldots, \mathbf{x}_p, \xi^1, \ldots, \xi^q) = T_\alpha^\beta x^\alpha \xi_\beta.$$

Cela démontre la

Proposition 1. *Tous les produits tensoriels de la forme*

$$\mathbf{e}^\alpha \otimes \mathbf{e}_\beta = \mathbf{e}^{i_1} \otimes \ldots \otimes \mathbf{e}^{i_p} \otimes \mathbf{e}_{j_1} \otimes \ldots \otimes \mathbf{e}_{j_q}$$

forment une base de l'espace $\mathsf{T}_p^q(\mathcal{V})$. *Les coordonnées d'un tenseur* T *dans cette base sont ses coefficients* :

$$T = T_{i_1 \ldots i_p}^{j_1 \ldots j_q} \mathbf{e}^{i_1} \otimes \ldots \otimes \mathbf{e}^{i_p} \otimes \mathbf{e}_{j_1} \otimes \ldots \otimes \mathbf{e}_{j_q} = T_\alpha^\beta \mathbf{e}^\alpha \otimes \mathbf{e}_\beta. \quad \square$$

Cette proposition relative aux fonctionnelles bilinéaires nous est connue dès la leçon 5.

En particulier,

$$\dim \mathsf{T}_p^q(\mathcal{V}) = n^{p+q},$$

si bien que la dimension de l'espace des tenseurs de type (p, q) est égale au nombre de leurs composantes, comme il fallait s'y attendre.

* * *

Etant donné un tenseur T quelconque de type (p, q), $p > 0$, $q > 0$, on suppose que $1 \leqslant k \leqslant p$, $1 \leqslant l \leqslant q$. On remplace la k-ième variable vectorielle par le vecteur de base $\mathbf{e}_i$ et la l-ième varia-

ble covectorielle par le covecteur $\mathbf{e}^i$ et on somme en i de 1 à n, il vient un tenseur S de type $(p-1, q-1)$. Ainsi,

$$S(\mathbf{x}_1, \ldots, \mathbf{x}_{p-1}, \xi^1, \ldots, \xi^{q-1}) =$$
$$= T(\mathbf{x}_1, \ldots, \mathbf{x}_{k-1}, \mathbf{e}_i, \mathbf{x}_k, \ldots, \mathbf{x}_{p-1}, \xi^1, \ldots, \xi^{l-1}, \mathbf{e}^i, \xi^l, \ldots, \xi^{q-1}),$$

la sommation par rapport à i étant sous-entendue à droite en vertu de la convention d'Einstein. Les composantes de S s'expriment manifestement par la formule

$$S^{j_1 \ldots j_{q-1}}_{i_1 \ldots i_{p-1}} = T^{j_1 \ldots j_{l-1} i j_l \ldots j_{q-1}}_{i_1 \ldots i_{k-1} i i_k \ldots i_{p-1}}.$$

Définition 1. On dit que le tenseur S est obtenu par *contraction relative au k-ième indice inférieur et au l-ième indice supérieur du tenseur T.*

On vérifie si la définition est intrinsèque, i.e. si S est indépendant de la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$, ce qui ne présente aucune difficulté. En effet, si $\mathbf{e}_{1'}, \ldots, \mathbf{e}_{n'}$ est toute autre base et si

$$\mathbf{e}_{i'} = c^i_{i'} \mathbf{e}_i,$$

on remplace les variables « non contractées» par les points de suspension et on a

$$T(\mathbf{x}_1, \ldots, \mathbf{x}_{k-1}, \mathbf{e}_{i'}, \mathbf{x}_k, \ldots, \mathbf{x}_{p-1}, \xi^1, \ldots, \xi^{l-1}, \mathbf{e}^{i'}, \xi^l, \ldots, \xi^{q-1}) =$$
$$= T(\ldots, \mathbf{e}_{i'}, \ldots, \mathbf{e}^{i'}, \ldots) = c^i_{i'} c^{i'}_j T(\ldots, \mathbf{e}_i, \ldots, \mathbf{e}^j, \ldots) =$$
$$= \delta^j_i T(\ldots, \mathbf{e}_i, \ldots, \mathbf{e}^j, \ldots) = T(\ldots, \mathbf{e}_i, \ldots, \mathbf{e}^j, \ldots) =$$
$$= S(\mathbf{x}_1, \ldots, \mathbf{x}_{p-1}, \xi^1, \ldots, \xi^{q-1}). \quad \square$$

Exemples de tenseurs contractés.

1. La contraction relative à i, j de la fonctionnelle bilinéaire mixte $B(\mathbf{x}, \xi) = b^j_i x^i \xi^j$ donne un tenseur de type (0, 0), i. e. le nombre $B(\mathbf{e}_i, \mathbf{e}^i)$. Ce nombre est appelé la *trace* de la fonctionnelle et noté tr B. Ainsi, on a par définition

$$\operatorname{tr} B = b^i_i = b^1_1 + \ldots + b^n_n,$$

d'où l'on tire que *la trace d'une fonctionnelle est égale à la trace de sa matrice*, i.e. à la somme des éléments diagonaux de celle-ci.

2. En particulier, on a pour tout vecteur $\mathbf{x}$ et tout covecteur ξ :

$$\operatorname{tr}(\xi \otimes \mathbf{x}) = \xi_i x^i = \xi_1 x^1 + \ldots \xi_n x^n.$$

3. Soit T un tenseur arbitraire de type (p, q). On prend p vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_p$ et q covecteurs $\xi^1, \ldots, \xi^q$ et on forme le tenseur

$$\mathbf{x}_1 \otimes \ldots \otimes \mathbf{x}_p \otimes T \otimes \xi^1 \otimes \ldots \otimes \xi^q$$

de type $(p+q, p+q)$. On le contracte $p+q$ fois relativement aux indices inférieurs et supérieurs de même numéro et on a évidem-

ment le nombre

$$T^{j_1 \ldots j_q}_{i_1 \ldots i_p} x_1^{i_1} \ldots x_p^{i_p} \xi^1_{j_1} \ldots \xi^q_{j_q},$$

i. e. la valeur de T sur les vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_p$ et les covecteurs $\xi^1, \ldots, \xi^q$.

* * *

Le cas le plus intéressant est celui des tenseurs de type $(p, 0)$ connus sous le nom de *fonctionnelles multilinéaires* sur l'espace vectoriel $\mathcal{V}$, dont le *degré* est par définition le nombre p de variables.

Etant donnée une base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de l'espace $\mathcal{V}$, chaque fonctionnelle multilinéaire A de degré p est définie de façon unique par ses coefficients:

$$A_{i_1 \ldots i_p} = A(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_p})$$

moyennant chacune de deux formules équivalentes

$$A(\mathbf{x}_1, \ldots, \mathbf{x}_p) = A_{i_1 \ldots i_p} x_1^{i_1} \ldots x_p^{i_p}$$

et

$$A = A_{i_1 \ldots i_p} \mathbf{e}^{i_1} \otimes \ldots \otimes \mathbf{e}^{i_p}.$$

Si l'on fixe toutes les variables de A, sauf une, on obtient une fonctionnelle de degré 1, i. e. un covecteur.

Définition 2. On dit de chacun de ces covecteurs qu'il est *associé* à la fonctionnelle multilinéaire A.

On obtient un covecteur associé ξ quelconque en se donnant $p - 1$ vecteurs $\mathbf{a}_1, \ldots, \mathbf{a}_{p-1}$ et le numéro i, auquel cas on définit ξ par

$$\xi(\mathbf{x}) = A(\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{x}, \mathbf{a}_i, \ldots, \mathbf{a}_{p-1}).$$

Définition 3. Le sous-espace de $\mathcal{V}'$ engendré par tous les covecteurs associés à une fonctionnelle multilinéaire A s'appelle l'*espace associé* à la fonctionnelle.

Définition 4. Si une fonctionnelle multilinéaire A de degré p est combinaison linéaire des produits tensoriels de la forme $\xi^{j_1} \otimes \ldots \otimes \xi^{j_p}$, où $1 \leqslant j_1, \ldots, j_p \leqslant r$, on dit qu'elle *s'exprime tensoriellement* au moyen des covecteurs $\xi^1, \ldots, \xi^r$.

Proposition 2. *Toute fonctionnelle multilinéaire A s'exprime tensoriellement par chaque base de l'espace associé.*

Démonstration. Soient $\mathcal{R}$ l'espace associé à A et $\mathbf{e}^1, \ldots, \mathbf{e}^r$ sa base quelconque. On complète $\mathbf{e}^1, \ldots, \mathbf{e}^r$ de façon à obtenir la base

$$\mathbf{e}^1, \ldots, \mathbf{e}^r, \ldots, \mathbf{e}^n$$

de l'espace $\mathcal{V}'$ tout entier et on considère la base duale

$$\mathbf{e}_1, \ldots, \mathbf{e}_r, \ldots, \mathbf{e}_n$$

de $\mathcal{T} = (\mathcal{T}')'$. On sait (voir leçon 4) que les vecteurs

$$\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n$$

constituent une base de l'annulateur $\mathcal{R}^\circ$ de $\mathcal{R}$, si bien que

$$\xi\ (\mathbf{e}_j) = 0 \text{ pour } j > r$$

quel que soit le covecteur ξ de $\mathcal{R}$. Mais parmi les éléments de $\mathcal{R}$ on trouve en particulier tous les covecteurs de la forme

$$\xi(\mathbf{x}) = A(\mathbf{x}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_p}).$$

Quels que soient les indices $i_2, \ldots, i_p$ et $i_1 > r$, on a donc l'égalité

$$A(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_p}) = 0,$$

i. e.

$$A_{i_1 i_2 \ldots i_p} = 0.$$

On démontre de même que cette égalité a lieu non seulement pour $i_1 > r$, mais aussi pour $i_2 > r$ et, en général, si $i_k > r$ pour au moins un $k = 1, 2, \ldots, p$. Cela étant, on estime qu'on somme dans

$$A = A_{i_1 \ldots i_p} \mathbf{e}^{i_1} \otimes \ldots \otimes \mathbf{e}^{i_p}$$

par rapport à tous les indices seulement de 1 à r, ce qui signifie que A s'exprime tensoriellement en fonction de la base $\mathbf{e}^1, \ldots, \mathbf{e}^r$. $\square$

LEÇON 7

Rang d'une fonctionnelle multilinéaire. — Fonctionnelles et permutations. — Application alternée.

On continue l'étude de l'espace associé à une fonctionnelle multilinéaire.

Soient A une fonctionnelle multilinéaire et $\mathscr{R}$ l'espace associé. Soit de plus $\xi^1, \ldots, \xi^r$ une famille arbitraire de covecteurs dont A est combinaison tensorielle.

Proposition 1. *Le sous-espace $\mathscr{R}$ est contenu dans l'enveloppe linéaire des covecteurs $\xi^1, \ldots, \xi^r$:*

$$\mathscr{R} \subset [\xi^1, \ldots, \xi^r].$$

Démonstration. On a par hypothèse

$$A = b_{i_1 \ldots i_p} \xi^{i_1} \otimes \ldots \otimes \xi^{i_p},$$

$b_{i_1 \ldots i_p}$ étant des nombres et la sommation étant étendue aux indices $i_1, \ldots, i_p$ de 1 à r.

Un covecteur quelconque

$$\xi(\mathbf{x}) = A(\mathbf{a}_1, \ldots, \mathbf{a}_{s-1}, \mathbf{x}, \mathbf{a}_s, \ldots, \mathbf{a}_{p-1}),$$

associé à la fonctionnelle A, est donc défini par la formule

$$\xi = c_q \xi^q,$$

où

$$c_q = b_{i_1 \ldots i_{s-1} q i_s \ldots i_{p-1}} a_1^{i_1} \ldots a_{s-1}^{i_{s-1}} a_s^{i_s} \ldots a_{p-1}^{i_{p-1}}.$$

Par conséquent, $\xi \in [\xi^1, \ldots, \xi^r]$, donc $\mathscr{R} \subset [\xi^1, \ldots, \xi^r]$.□

Définition 1. La dimension de l'espace $\mathscr{R}$ constitue le rang de la fonctionnelle multilinéaire A.

Théorème 1. *Le rang de la fonctionnelle multilinéaire A est égal au plus petit nombre de covecteurs dont A est combinaison tensorielle, i. e.*

a) *si A s'exprime tensoriellement en fonction des covecteurs $\xi^1, \ldots, \xi^r$, son rang est au plus égal à r ;*

b) *si r est le rang de la fonctionnelle A, on trouve r covecteurs $\xi^1, \ldots, \xi^r$ qui représentent tensoriellement A.*

Cela étant, la famille $\xi^1, \ldots, \xi^r$ *jouit de la propriété* b) *si et seulement si c'est une base de l'espace* $\mathscr{R}$ *associé à* A.

Démonstration. Aux termes de la proposition 2 de la leçon précédente, la fonctionnelle A s'exprime tensoriellement par une base de $\mathscr{R}$, ce qui démontre en particulier l'affirmation b).

Si A s'exprime tensoriellement au moyen des covecteurs $\xi^1, \ldots, \xi^r$ et on a donc l'inclusion

$$\mathscr{R} \subset [\xi^1, \ldots, \xi^r]$$

(proposition 1), alors

$$\dim \mathscr{R} \leqslant \dim [\xi^1, \ldots, \xi^r] \leqslant r,$$

ce qui démontre a).

On voit de plus qu'on a nécessairement pour $r = \dim \mathscr{R}$:

$$\mathscr{R} = [\xi^1, \ldots, \xi^r],$$

i. e. la famille $\xi^1, \ldots, \xi^r$ (qui est visiblement libre) est une base de $\mathscr{R}$.

Le théorème se trouve démontré. □

On conçoit que ces résultats restent vrais (à des modifications évidentes près) pour les fonctionnelles des *covecteurs* (i. e. pour les tenseurs de type $(0, p)$). On n'oubliera pas que les éléments associés à ces fonctionnelles sont des *vecteurs*, si bien que l'espace associé est un sous-espace de $\mathscr{T}$.

* * *

On rappelle qu'une *permutation de degré* p est par définition une bijection quelconque de l'ensemble $\{1, \ldots, p\}$ sur lui-même. Toute permutation σ est en général représentée par le tableau à deux lignes

$$\begin{pmatrix} 1 & 2 & \ldots & p \\ \sigma(1) & \sigma(2) & \ldots & \sigma(p) \end{pmatrix}$$

bien que la ligne inférieure suffise en général à elle seule.

Toutes les permutations de degré p forment un groupe (pour la composition). Ce groupe est appelé *groupe symétrique* et noté S_p.

Une permutation est *paire* ou *impaire* selon que les couples $(\sigma(i), \sigma(j))$, $i < j$ et $\sigma(i) > \sigma(j)$ sont en nombre pair ou impair.

Etant donnée une permutation σ, on appelle *signature* de σ le nombre $\varepsilon_\sigma = \pm 1$, le signe plus correspondant à σ paire et le signe moins à σ impaire.

Il est connu que

$$\varepsilon_{\sigma\tau} = \varepsilon_\sigma \varepsilon_\tau$$

pour deux permutations σ et τ quelconques, d'où en particulier *la propriété, pour toutes les permutations paires, de former un sous-groupe de* S_p.

Soit A une fonctionnelle multilinéaire de degré p.

Définition 2. Soit une permutation $\sigma \in S_p$ quelconque. Le symbole σA désigne une fonctionnelle définie par la formule

$$(\sigma A)(x_1, \ldots, x_p) = A(x_{\sigma(1)}, \ldots, x_{\sigma(p)}).$$

Il est clair que

$$(\sigma A)_{i_1 \ldots i_p} = A_{i_{\sigma(1)} \ldots i_{\sigma(p)}}.$$

Ainsi, *on obtient les coefficients de la fonctionnelle σA par permutation σ des indices des coefficients de A.*

Exemple. Soit $n = 5$, $p = 3$ et

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

On a

$$(\sigma A)_{145} = A_{514}, \quad (\sigma A)_{553} = A_{355}.$$

L'application

$$A \mapsto \sigma A$$

est évidemment, pour toute permutation σ, une application linéaire (un homomorphisme) de l'espace vectoriel $\mathsf{T}_p(\mathcal{V})$ sur lui-même. On constate en outre sans peine que

$$(\sigma\tau) A = \sigma(\tau A)$$

pour $\sigma, \tau \in S_p$ quelconques, ce qui entraîne en particulier que *l'application $A \mapsto \sigma A$ est un isomorphisme.* □

* * *

Dans la suite de la leçon, nous supposerons que *le corps de base* K *est de caractéristique* 0, i. e. on définit dans K la division par tout entier naturel (et en particulier par la factorielle p !).

Définition 3. Etant donnée une fonctionnelle $A \in \mathsf{T}_p(\mathcal{V})$ quelconque, on désigne par Alt A la fonctionnelle

$$\mathrm{Alt}\, A = \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma (\sigma A).$$

Il est clair que l'application

$$A \mapsto \mathrm{Alt}\, A$$

est linéaire (c'est un homomorphisme). Elle est dite *alternée.*

Puisque $\mathrm{Alt}: \mathsf{T}_p(\mathcal{V}) \to \mathsf{T}_p(\mathcal{V})$ et $\sigma: \mathsf{T}_p(\mathcal{V}) \to \mathsf{T}_p(\mathcal{V})$, on définit les applications composées $\sigma \circ \mathrm{Alt}$ et $\mathrm{Alt} \circ \sigma$.

Proposition 2. *Pour toute permutation $\sigma \in S_p$, on a les relations*

$$\mathrm{Alt} \circ \sigma = \varepsilon_\sigma \mathrm{Alt}, \quad \sigma \circ \mathrm{Alt} = \varepsilon_\sigma \mathrm{Alt}.$$

Démonstration. Toute fonctionnelle $A \in \mathsf{T}_p(\mathcal{F})$ satisfait à

$$\operatorname{Alt}(\sigma A) = \frac{1}{p!} \sum_{\tau \in S_p} \varepsilon_\tau (\tau\sigma A) =$$

$$= \varepsilon_\sigma \frac{1}{p!} \sum_{\tau \in S_p} \varepsilon_{\tau\sigma} (\tau\sigma A) = \varepsilon_\sigma \frac{1}{p!} \sum_{\tau \in S_p} \varepsilon_\tau (\tau A) = \varepsilon_\sigma \operatorname{Alt} A,$$

car $\tau\sigma$ parcourt S_p tout entier avec τ.

De même,

$$\sigma \operatorname{Alt} A = \frac{1}{p!} \sum_{\tau \in S_p} \varepsilon_\tau (\sigma\tau A) = \varepsilon_\sigma \frac{1}{p!} \sum_{\tau \in S_p} \varepsilon_{\sigma\tau} (\sigma\tau A) = \varepsilon_\sigma \operatorname{Alt} A. \quad \square$$

Puisque $\operatorname{Alt}: \mathsf{T}_p(\mathcal{F}) \to \mathsf{T}_p(\mathcal{F})$, on définit l'itération

$$\operatorname{Alt} \circ \operatorname{Alt}: \mathsf{T}_p(\mathcal{F}) \to \mathsf{T}_p(\mathcal{F}).$$

Proposition 3. *On a l'égalité*

$$\operatorname{Alt} \circ \operatorname{Alt} = \operatorname{Alt}.$$

Démonstration. Une application alternée étant linéaire, on a par suite de la proposition 2

$$\operatorname{Alt}(\operatorname{Alt} A) = \operatorname{Alt}\left(\frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma (\sigma A)\right) =$$

$$= \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma \operatorname{Alt}(\sigma A) = \frac{1}{p!} \sum_{\sigma \in S_p} (\varepsilon_\sigma)^2 \operatorname{Alt} A = \operatorname{Alt} A. \quad \square$$

Comment exprimer les coefficients $(\operatorname{Alt} A)_{i_1 \ldots i_p}$ de la fonctionnelle $\operatorname{Alt} A$ au moyen des coefficients $A_{i_1 \ldots i_p}$ de A ? Il y a intérêt, pour avoir des formules condensées, à introduire une notation qui rendra également d'autres services.

Etant donnés n^p nombres $A_{i_1 \ldots i_p}$ numérotés par les indices $i_1, \ldots, i_p$ variant de 1 à n, on fait leur correspondre n^p nombres $B_{i_1 \ldots i_p}$ numérotés de façon analogue tels que

$$B_{i_1 \ldots i_p} = \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma A_{i_{\sigma(1)} \ldots i_{\sigma(p)}}.$$

On désigne d'ordinaire les nombres $B_{i_1 \ldots i_p}$ par $A_{[i_1 \ldots i_p]}$, ce qui nuit d'une certaine façon à la rigueur des formules. Ainsi, on a par définition

$$A_{[i_1 \ldots i_p]} = \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma A_{i_{\sigma(1)} \ldots i_{\sigma(p)}}.$$

Avec ces notations, la position des indices est certes indifférente. Si n^p nombres donnés s'écrivent $A^{i_1 \dots i_p}$, alors on pose

$$A^{[i_1 \dots i_p]} = \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma A^{i_{\sigma(1)} \dots i_{\sigma(p)}}.$$

Proposition 4. *Les coefficients de la fonctionnelle* Alt A *vérifient la formule*

$$(\mathrm{Alt}\, A)_{i_1 \dots i_p} = A_{[i_1 \dots i_p]}.$$

Démonstration. Par définition

$$(\mathrm{Alt}\, A)_{i_1 \dots i_p} = \mathrm{Alt}\, A\, (\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_p}) =$$

$$= \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma (\sigma A)(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_p}) = \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma A\, (\mathbf{e}_{i_{\sigma(1)}}, \dots, \mathbf{e}_{i_{\sigma(p)}}) =$$

$$= \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma A_{i_{\sigma(1)} \dots i_{\sigma(p)}} = A_{[i_1 \dots i_p]}. \quad \square$$

L'intérêt des notations introduites plus haut ne se limite pas à cette formule. On s'en sert par exemple en théorie des déterminants.

Lemme 1. *On a l'identité*

$$\begin{vmatrix} x_1^1 & \dots & x_1^p \\ x_2^1 & \dots & x_2^p \\ \cdot & \cdot & \cdot \\ x_p^1 & \dots & x_p^p \end{vmatrix} = p!\, x_1^{[1} \dots x_p^{p]} = p!\, x_{[1}^1 \dots x_{p]}^p.$$

Démonstration. Par définition

$$p!\, x_1^{[1} \dots x_p^{p]} = \sum_{\sigma \in S_p} \varepsilon_\sigma x_1^{\sigma(1)} \dots x_p^{\sigma(p)}$$

et

$$p!\, x_{[1}^1 \dots x_{p]}^p = \sum_{\sigma \in S_p} \varepsilon_\sigma x_{\sigma(1)}^1 \dots x_{\sigma(p)}^p.$$

Les deux seconds membres sont égaux au déterminant $|x_i^j|$ (qui est développé « suivant les lignes » ou « suivant les colonnes »). $\square$

Proposition 5. *Quels que soient les vecteurs* $\mathbf{x}_1, \dots, \mathbf{x}_p$, *on a les formules*

$$(\mathrm{Alt}\, A)(\mathbf{x}_1, \dots \mathbf{x}_p) = A_{[i_1 \dots i_p]} x_1^{i_1} \dots x_p^{i_p} =$$

$$= A_{i_1 \dots i_p} x_{[1}^{i_1} \dots x_{p]}^{i_p} = A_{i_1 \dots i_p} x_1^{[i_1} \dots x_p^{i_p]}.$$

Démonstration. La première formule n'est rien autre que la proposition 4 écrite de façon différente, la deuxième est établie par

les calculs

$$(\mathrm{Alt}\,A)(\mathbf{x}_1, \ldots, \mathbf{x}_p) = \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma A(x_{\sigma(1)}, \ldots, x_{\sigma(p)}) =$$

$$= \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma A_{i_1 \ldots i_p} x^{i_1}_{\sigma(1)} \ldots x^{i_p}_{\sigma(p)} =$$

$$= A_{i_1 \ldots i_p} \left(\frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma x^{i_1}_{\sigma(1)} \ldots x^{i_p}_{\sigma(p)} \right) =$$

$$= A_{i_1 \ldots i_p} x^{i_1}_{[1} \ldots x^{i_p}_{p]},$$

et la troisième découle de la deuxième par suite du lemme 1. □

Conséquence. *On a la formule*

$$(\mathrm{Alt}\,A)(\mathbf{x}_1, \ldots, \mathbf{x}_p) = \frac{1}{p!} A_{i_1 \ldots i_p} \begin{vmatrix} x^{i_1}_1 & \ldots & x^{i_1}_p \\ x^{i_2}_1 & \ldots & x^{i_2}_p \\ \cdot & \cdot & \cdot \\ x^{i_p}_1 & \ldots & x^{i_p}_p \end{vmatrix},$$

où l'on somme comme toujours par rapport à $i_1, \ldots, i_p$ de 1 à n.

Exemple. Quand $p = 2$,

$$(\mathrm{Alt}\,A)(\mathbf{x}, \mathbf{y}) = \left(\frac{A_{ij} - A_{ji}}{2} \right) x^i y^j = A_{ij} \left(\frac{x^i y^j - y^i x^j}{2} \right) =$$

$$= A_{ij} \left(\frac{x^i y^j - x^j y^i}{2} \right) = \frac{1}{2} A_{ij} \begin{vmatrix} x^i & y^i \\ x^j & y^j \end{vmatrix}.$$

LEÇON 8

Fonctionnelles multilinéaires antisymétriques. — Produit extérieur. — Algèbre de Grassmann. — Produit extérieur de covecteurs. — Décomposition des fonctionnelles antisymétriques en produit extérieur de covecteurs de base.

Soit donnés deux ensembles de n^p nombres $x_{i_1 \ldots i_p}$ et $y^{i_1 \ldots i_p}$ numérotés par p indices $i_1, \ldots, i_p$ qui parcourent de façon indépendante les valeurs de 1 à n. On multiplie chaque $x_{i_1 \ldots i_p}$ par $y^{i_1 \ldots i_p}$ correspondant et on fait la somme, il vient le nombre

$$x_{i_1 \ldots i_p} y^{i_1 \ldots i_p}.$$

Lemme 1. *Pour toute permutation σ du groupe symétrique S_p on a l'identité*

$$x_{i_1 \ldots i_p} y^{i_1 \ldots i_p} = x_{i_{\sigma(1)} \ldots i_{\sigma(p)}} y^{i_{\sigma(1)} \ldots i_{\sigma(p)}}. \tag{1}$$

Démonstration. Les deux membres de (1) sont les sommes de mêmes termes à l'ordre près. □

Soit n^2 nombres x^i_j, avec $i, j = 1, \ldots, n$. Considérons tous les produits

$$x^{i_1}_{j_1} \ldots x^{i_p}_{j_p}.$$

Lemme 2. *Pour toute permutation $\sigma \in S_p$, on a l'identité*

$$x^{i_{\sigma(1)}}_{j_1} \ldots x^{i_{\sigma(p)}}_{j_p} = x^{i_1}_{j_{\tau(1)}} \ldots x^{i_p}_{j_{\tau(p)}}, \tag{2}$$

où $\tau = \sigma^{-1}$.

Démonstration. Les deux membres de (2) sont le produit de mêmes facteurs à l'ordre près. Si l'on a par exemple $p = 4$ et

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix},$$

alors

$$x^{i_2}_{j_1} x^{i_4}_{j_2} x^{i_3}_{j_3} x^{i_1}_{j_4} = x^{i_1}_{j_4} x^{i_2}_{j_1} x^{i_3}_{j_3} x^{i_4}_{j_2}. \quad \square$$

Définition 1. Une fonctionnelle multilinéaire A est dite *antisymétrique* si

$$\sigma A = \varepsilon_\sigma A$$

pour toute permutation $\sigma \in S_p$.

Proposition 1. *Une fonctionnelle A est antisymétrique si et seulement si*

$$A_{i_{\sigma(1)}\ldots i_{\sigma(p)}} = \varepsilon_\sigma A_{i_1\ldots i_p} \tag{3}$$

quels que soient les indices $i_1, \ldots, i_p$ et la permutation $\sigma \in S_p$.

Démonstration. Si A est antisymétrique, alors

$$A_{i_{\sigma(1)}\ldots i_{\sigma(p)}} = (\sigma A)_{i_1\ldots i_p} = \varepsilon_\sigma A_{i_1\ldots i_p}.$$

Inversement, si l'on a (3), alors tous les vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_n$ vérifient les égalités

$$(\sigma A)(\mathbf{x}_1 \ldots, \mathbf{x}_p) = A(\mathbf{x}_{\sigma(1)}, \ldots, \mathbf{x}_{\sigma(p)}) = A_{i_1\ldots i_p} x^{i_1}_{\sigma(1)} \ldots x^{i_p}_{\sigma(p)}.$$

Or, on a par les lemmes 1 et 2 et la condition (3):

$$A_{i_1\ldots i_p} x^{i_1}_{\sigma(1)} \ldots x^{i_p}_{\sigma(p)} = A_{i_{\sigma(1)}\ldots i_{\sigma(p)}} x^{i_{\sigma(1)}}_{\sigma(1)} \ldots x^{i_{\sigma(p)}}_{\sigma(p)} =$$

$$= A_{i_{\sigma(1)}\ldots i_{\sigma(p)}} x^{i_1}_1 \ldots x^{i_p}_p = \varepsilon_\sigma A_{i_1\ldots i_p} x^{i_1}_1 \ldots x^{i_p}_p = \varepsilon_\sigma A(\mathbf{x}_1, \ldots \mathbf{x}_p).$$

Donc, $\sigma A = \varepsilon_\sigma A$. □

Proposition 2. *Une fonctionnelle multilinéaire A est antisymétrique si et seulement si elle est conservée par toute application alternée, i.e. si*

$$\operatorname{Alt} A = A.$$

Démonstration. Si

$$\sigma A = \varepsilon_\sigma A$$

pour toute permutation $\sigma \in S_p$, alors tous les termes de la somme

$$\sum_{\sigma \in S_p} \varepsilon_\sigma (\sigma A)$$

sont égaux à A, si bien que la somme vaut $p!A$. Par conséquent,

$$\operatorname{Alt} A = A.$$

Inversement, si $\operatorname{Alt} A = A$, on a, d'après la proposition 2 de la leçon 7,

$$\sigma A = \sigma \operatorname{Alt} A = \varepsilon_\sigma \operatorname{Alt} A = \varepsilon_\sigma A. \ \square$$

Conséquence. *Une fonctionnelle multilinéaire A est antisymétrique si et seulement si ses coefficients satisfont aux égalités*

$$A_{i_1\ldots i_p} = A_{[i_1\ldots i_p]}.$$

Une condition un peu plus générale pour qu'une fonctionnelle jouisse de la propriété étudiée est fournie formellement par la

Proposition 3. *Une fonctionnelle multilinéaire A est antisymétrique si et seulement s'il existe une fonctionnelle multilinéaire B telle que*

$$A = \operatorname{Alt} B. \tag{4}$$

Démonstration. Si A est antisymétrique, alors (4) a lieu pour $B = A$ (proposition 2). Inversement, si (4) est vérifiée, alors

$$\mathrm{Alt}\, A = \mathrm{Alt}\,(\mathrm{Alt}\, B) = \mathrm{Alt}\, B = A$$

(leçon 7, prop. 3), si bien que A est antisymétrique (proposition 2). □

* * *

Le produit tensoriel $A \otimes B$ de deux fonctionnelles antisymétriques A et B n'est en général pas antisymétrique. Pour que $A \otimes B$ le soit, il faut lui faire subir une application alternée.

Définition 2. Etant données deux fonctionnelles multilinéaires antisymétriques A et B, on appelle leur *produit extérieur* la fonctionnelle

$$A \wedge B = \mathrm{Alt}\,(A \otimes B).$$

La fonctionnelle $A \wedge B$ est de degré $p + q$, p et q étant les degrés de A et B respectivement, et ses coefficients sont définis par

$$(A \wedge B)_{i_1 \ldots i_{p+q}} = A_{[i_1 \ldots i_p} B_{i_{p+1} \ldots i_{p+q}]}.$$

Proposition 4. *La multiplication extérieure de fonctionnelles antisymétrique est associative, i.e.*

$$(A \wedge B) \wedge C = A \wedge (B \wedge C)$$

quelles que soient les fonctionnelles antisymétriques A, B et C.

On omet donc les parenthèses dans le produit extérieur de plusieurs fonctionnelles.

Avant de démontrer la proposition 4, nous tenons à faire plusieurs remarques qui ont également leur intérêt propre.

Quels que soient p et q, on applique le groupe symétrique S_p dans le groupe symétrique S_{p+q} en associant à toute permutation $\sigma \in S_p$ une permutation $\sigma' \in S_{p+q}$ qui opère sur $1, \ldots, p$ comme σ et qui laisse invariants les nombres $p + 1, \ldots, q$:

$$\sigma'(i) = \begin{cases} \sigma(i) & \text{si } 1 \leqslant i \leqslant p, \\ i & \text{si } p+1 \leqslant i \leqslant p+q. \end{cases}$$

Il est clair que la correspondance $\sigma \mapsto \sigma'$ est un monomorphisme (un homomorphisme injectif) conservant la signature, i.e.

$$\varepsilon_{\sigma'} = \varepsilon_\sigma$$

pour toute permutation σ.

Les permutations $\sigma \in S_p$ permettent d'effectuer, pour une fonctionnelle multilinéaire A quelconque de degré $p + q$, une application alternée par rapport aux p premières variables, i.e. de construire

la fonctionnelle

$$\operatorname{alt} A = \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma (\sigma' A).$$

Lemme 3.

$$\operatorname{Alt}(\operatorname{alt} A) = \operatorname{Alt} A.$$

Démonstration. On répète les raisonnements de la proposition 3 de la leçon précédente:

$$\operatorname{Alt}(\operatorname{alt} A) = \operatorname{Alt}\left(\frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma (\sigma' A)\right) =$$

$$= \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma \varepsilon_{\sigma^{-1}} \operatorname{Alt} A = \operatorname{Alt} A. \quad \square$$

Démonstration de la proposition 4. Soient p, q et r les degrés respectifs des fonctionnelles A, B et C.

Par définition,

$$(A \wedge B) \wedge C = \operatorname{Alt}((A \wedge B) \otimes C).$$

Mais

$$(A \wedge B) \otimes C = \operatorname{alt}(A \otimes B \otimes C),$$

où alt désigne l'application alternée relative aux $p + q$ premiers indices. On a donc, d'après le lemme 3,

$$(A \wedge B) \wedge C = \operatorname{Alt}(A \otimes B \otimes C).$$

On démontre de même l'égalité

$$A \wedge (B \wedge C) = \operatorname{Alt}(A \otimes B \otimes C).$$

Par conséquent, $(A \wedge B) \wedge C = A \wedge (B \wedge C)$. $\square$

En particulier,

$$A \wedge B \wedge C = \operatorname{Alt}(A \otimes B \otimes C).$$

Une formule analogue a évidemment lieu pour les facteurs en nombre arbitraire.

A la différence du produit tensoriel, le produit extérieur est commutatif (encore qu'au signe près).

Proposition 5. *Etant données deux fonctionnelles multilinéaires antisymétriques A et B quelconques de degré p et q respectivement, on a l'égalité*

$$B \wedge A = (-1)^{pq} A \wedge B.$$

Cette propriété du produit extérieur s'appelle l'*anticommutativité*.

Démonstration. On a par définition

$$\begin{aligned}(B \otimes A)(\mathbf{x}_1, \ldots, \mathbf{x}_{p+q}) &= B(\mathbf{x}_1, \ldots, \mathbf{x}_q)\, A(\mathbf{x}_{q+1}, \ldots, \mathbf{x}_{p+q}) = \\ &= A(\mathbf{x}_{q+1}, \ldots, \mathbf{x}_{p+q})\, B(\mathbf{x}_1, \ldots, \mathbf{x}_q) = \\ &= (A \otimes B)(\mathbf{x}_{q+1}, \ldots, \mathbf{x}_{p+q}, \mathbf{x}_1, \ldots, \mathbf{x}_q) = \\ &= (\sigma_0 (A \otimes B))(\mathbf{x}_1, \ldots, \mathbf{x}_{p+q}),\end{aligned}$$

où

$$\sigma_0 = \begin{pmatrix} 1, & \ldots, & p, & p+1, & \ldots, & p+q \\ q+1, & \ldots, & p+q, & 1, & \ldots, & q \end{pmatrix},$$

i.e.

$$B \otimes A = \sigma_0 (A \otimes B).$$

Aussi

$$B \wedge A = \mathrm{Alt}(B \otimes A) = \varepsilon_{\sigma_0} \mathrm{Alt}(A \otimes B) = \varepsilon_{\sigma_0} (A \wedge B).$$

La démonstration s'achève par l'égalité

$$\varepsilon_{\sigma_0} = (-1)^{pq}. \quad \square$$

* * *

Il est clair que l'ensemble $\Lambda_p(\mathscr{V})$ de toutes les fonctionnelles antisymétriques de degré p constitue un sous-espace de l'espace $\mathsf{T}_p(\mathscr{V})$, si bien que $\Lambda_p(\mathscr{V})$ est un espace vectoriel. Cela étant, la multiplication extérieure des fonctionnelles antisymétriques est évidemment distributive par rapport à l'addition :

$$(A + B) \wedge C = A \wedge C + B \wedge C,$$

i.e. les espaces vectoriels

$$\Lambda_0(\mathscr{V}),\ \Lambda_1(\mathscr{V}),\ \ldots,\ \Lambda_p(\mathscr{V}),\ \ldots$$

sont pour les opérations $+$ et $\wedge$ un objet algébrique qui est un exemple d'*algèbre graduée*. Cette algèbre s'appelle *algèbre extérieure* (ou *algèbre de Grassmann*) de l'espace $\mathscr{V}$ et se note $\Lambda(\mathscr{V})$.

On observe que si $p = 1$, l'antisymétrie n'impose aucune restriction. Par conséquent,

$$\Lambda_1(\mathscr{V}) = \mathsf{T}_1(\mathscr{V}) = \mathscr{V}'.$$

Des considérations analogues conduisent aux égalités

$$\Lambda_0(\mathscr{V}) = \mathsf{T}_0(\mathscr{V}) = \mathbb{K}.$$

L'antisymétrie entraîne évidemment $A(\mathbf{x}_1, \ldots, \mathbf{x}_p) = 0$ si au moins deux des vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_p$ coïncident (cf. le cas analogue de déterminants). Comme les fonctionnelles sont multilinéaires, on a $A(\mathbf{x}_1, \ldots, \mathbf{x}_p) = 0$ *si au moins l'un des vecteurs* $\mathbf{x}_1, \ldots, \mathbf{x}_p$ *est com-*

binaison linéaire des autres. On l'a nécessairement pour $p > n$, si bien que

$$\Lambda_p(\mathcal{T}) = 0, \qquad p > n.$$

* * *

Le cas le plus intéressant est celui des produits extérieurs de fonctionnelles de degré 1, i.e. de covecteurs.

Selon la remarque ci-dessus,

$$\xi^1 \wedge \ldots \wedge \xi^p = \mathrm{Alt}\,(\xi^1 \otimes \ldots \otimes \xi^p)$$

quels que soient les covecteurs $\xi^1, \ldots, \xi^p$. Cela signifie qu'on a pour $\mathbf{x}_1, \ldots, \mathbf{x}_p$ quelconques

$$(\xi^1 \wedge \ldots \wedge \xi^p)(\mathbf{x}_1, \ldots, \mathbf{x}_p) =$$
$$= \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma (\xi^1 \otimes \ldots \otimes \xi^p)(\mathbf{x}_{\sigma(1)}, \ldots, \mathbf{x}_{\sigma(p)}) =$$
$$= \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma \xi^1(\mathbf{x}_{\sigma(1)}) \ldots \xi^p(\mathbf{x}_{\sigma(p)}),$$

i.e.

$$(5) \qquad (\xi^1 \wedge \ldots \wedge \xi^p)(\mathbf{x}_1, \ldots, \mathbf{x}_p) = \frac{1}{p!} \begin{vmatrix} \xi^1(\mathbf{x}_1) & \ldots & \xi^1(\mathbf{x}_p) \\ \cdot & \cdot & \cdot \\ \xi^p(\mathbf{x}_1) & \ldots & \xi^p(\mathbf{x}_p) \end{vmatrix}.$$

On écrit cette identité très importante de façon équivalente (voir lemme 1 de la leçon 7):

$$(\xi^1 \wedge \ldots \wedge \xi^p)(\mathbf{x}_1, \ldots, \mathbf{x}_p) = \xi^1(\mathbf{x}_{[1]}) \ldots \xi^p(\mathbf{x}_{p]}),$$

ou encore

$$(6) \qquad (\xi^1 \wedge \ldots \wedge \xi^p)(\mathbf{x}_1, \ldots, \mathbf{x}_p) = \xi^{[1}(\mathbf{x}_1) \ldots \xi^{p]}(\mathbf{x}_p).$$

Soit la fonctionnelle

$$\xi^{[1} \otimes \ldots \otimes \xi^{p]} = \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma \xi^{\sigma(1)} \otimes \ldots \otimes \xi^{\sigma(p)}.$$

Ses valeurs sur $\mathbf{x}_1, \ldots, \mathbf{x}_p$ sont fournies par la formule

$$(7) \qquad (\xi^{[1} \otimes \ldots \otimes \xi^{p]})(\mathbf{x}_1, \ldots, \mathbf{x}_p) =$$
$$= \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma (\xi^{\sigma(1)} \otimes \ldots \otimes \xi^{\sigma(p)})(\mathbf{x}_1, \ldots, \mathbf{x}_p) =$$
$$= \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon_\sigma \xi^{\sigma(1)}(\mathbf{x}_1) \ldots \xi^{\sigma(p)}(\mathbf{x}_p) = \xi^{[1}(\mathbf{x}_1) \ldots \xi^{p]}(\mathbf{x}_p).$$

Les formules (6) et (7) donnent lieu à la

Proposition 6. *Etant donnés des covecteurs* $\xi^1, \ldots, \xi^p$ *quelconques, on a l'égalité*

$$\xi^1 \wedge \ldots \wedge \xi^p = \xi^{[1} \otimes \ldots \otimes \xi^{p]}. \ \square$$

Conséquence. *La fonctionnelle* $\xi^1 \wedge \ldots \wedge \xi^p$ *s'exprime tensoriellement par les covecteurs* $\xi^1, \ldots, \xi^p$. $\square$

Le résultat suivant est simple, mais important.

Proposition 7. *L'égalité*

$$\xi^1 \wedge \ldots \wedge \xi^p = 0$$

a lieu si et seulement s'il y a dépendance linéaire des covecteurs $\xi^1, \ldots, \xi^p$.

Démonstration. L'anticommutativité du produit extérieur fait que $\xi^1 \wedge \ldots \wedge \xi^p$ change de signe par échange de deux facteurs quelconques. On en déduit par des raisonnements familiers $\xi^1 \wedge \ldots \wedge \xi^p = 0$ si les covecteurs $\xi^1, \ldots, \xi^p$ sont linéairement dépendants.

Supposons qu'il y a indépendance linéaire des covecteurs, auquel cas on les complète de façon à obtenir la base

$$\mathbf{e}^1 = \xi^1, \ldots, \mathbf{e}^p = \xi^p, \quad \mathbf{e}^{p+1}, \ldots, \mathbf{e}^n$$

de $\mathscr{V}'$ tout entier. Soit

$$\mathbf{e}_1, \ldots, \mathbf{e}_n$$

la base duale de $\mathscr{V}$. On a conformément à (5):

$$(\xi^1 \wedge \ldots \wedge \xi^p)(\mathbf{e}_1, \ldots, \mathbf{e}_p) = \frac{1}{p!} \begin{vmatrix} \xi^1(\mathbf{e}_1) & \ldots & \xi^1(\mathbf{e}_p) \\ \cdot & \cdot & \cdot \\ \xi^p(\mathbf{e}_1) & \ldots & \xi^p(\mathbf{e}_p) \end{vmatrix} =$$

$$= \frac{1}{p!} \begin{vmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{vmatrix} = \frac{1}{p!} \neq 0.$$

Donc, $\xi^1 \wedge \ldots \wedge \xi^p \neq 0$. $\square$

* * *

Soit $\mathbf{e}^1, \ldots, \mathbf{e}^n$ une base arbitraire de $\mathscr{V}'$. Chaque fonctionnelle multilinéaire A admet (c'est connu) une représentation de la forme

$$A = A_{i_1 \ldots i_p} \mathbf{e}^{i_1} \otimes \ldots \otimes \mathbf{e}^{i_p}.$$

Si A est antisymétrique (donc $A = \operatorname{Alt} A$), on obtient par application alternée la formule

$$A = A_{i_1 \ldots i_p} \mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_p}. \tag{8}$$

De nombreux termes sont nuls et identiques. On réduit donc « les termes semblables ».

Selon la proposition 7, les termes à $i_1, \ldots, i_p$ non simultanément distincts sont nuls. Aussi

$$A = \sum A_{i_1 \ldots i_p} \mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_p}, \tag{9}$$

la sommation étant étendue de 1 à n à tous les groupes de p entiers $i_1, i_2, \ldots, i_p$ choisis parmi $1, 2, \ldots, p$ *sans répétition.*

On fixe un groupe et on considère les termes de (9) qui ne diffèrent que par l'ordre des indices. Ces termes sont au nombre de $p\,!$, et ils sont tous de la forme

$$\underbrace{A_{i_{\sigma(1)} \ldots i_{\sigma(p)}}}_{\text{(sans sommation)}} \mathbf{e}^{i_{\sigma(1)}} \wedge \ldots \wedge \mathbf{e}^{i_{\sigma(p)}}, \tag{10}$$

σ étant une permutation quelconque de degré p. L'anticommutativité du produit extérieur entraîne de suite

$$\mathbf{e}^{i_{\sigma(1)}} \wedge \ldots \wedge \mathbf{e}^{i_{\sigma(p)}} = \varepsilon_\sigma \mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_p}.$$

D'autre part, on a par la proposition 1

$$A_{i_{\sigma(1)} \ldots i_{\sigma(p)}} = \varepsilon_\sigma A_{i_1 \ldots i_p}.$$

Comme $\varepsilon_\sigma \varepsilon_\sigma = 1$, tous les termes (10) sont égaux à

$$\underbrace{A_{i_1 \ldots i_p}}_{\text{(sans sommation)}} \mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_p}.$$

On vient de démontrer l'égalité

$$A = p\,! \sum_{(i_1, \ldots, i_p)} A_{i_1 \ldots i_p} \mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_p},$$

la sommation étant étendue à toutes les *combinaisons* $(i_1, \ldots, i_p)$ d'indices de (9). Il existe pour toute combinaison un seul groupe $i_1, \ldots, i_p$ tel que $i_1 < i_2 < \ldots < i_p$, si bien qu'on a la

Proposition 8. *Pour toute fonctionnelle antisymétrique A on a l'égalité*

$$A = p\,! \sum_{i_1 < \ldots < i_p} A_{i_1 \ldots i_p} \mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_p}. \quad \square$$

Ainsi, les fonctionnelles du type

$$\mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_p}, \quad 1 \leqslant i_1 < \ldots < i_p \leqslant n,$$

constituent une famille génératrice dans $\wedge_p(\mathscr{V})$.

LEÇON 9

Base de l'espace des fonctionnelles antisymétriques. — Formules de transformation de la base de l'espace des fonctionnelles antisymétriques. — Multivecteurs. — Rang extérieur d'une fonctionnelle antisymétrique. — Théorème du rang d'un multivecteur. — Conditions d'égalité de deux multivecteurs.

La proposition 1 de la leçon précédente entraîne, pour les coefficients $A_{i_1 \ldots i_p}$ d'une fonctionnelle antisymétrique A, les égalités

$$(1) \qquad A_{i_1 \ldots i_p} = \begin{cases} 0 & \text{si les nombres } i_1, \ldots, i_p \text{ sont non tous distincts,} \\ \varepsilon_\sigma A_{i_{\sigma(1)} \ldots i_{\sigma(p)}} & \text{dans le cas contraire,} \end{cases}$$

avec σ une permutation de degré p telle que

$$i_{\sigma(1)} < \ldots < i_{\sigma(p)}.$$

Il en résulte qu'on définit complètement la fonctionnelle A connaissant $A_{i_1 \ldots i_p}$ pour lesquels $i_1 < \ldots < i_p$.

Définition 1. Les coefficients $A_{i_1 \ldots i_p}$ tels que $i_1 < \ldots < i_p$ sont appelés *coefficients stricts* de la fonctionnelle antisymétrique A.

Proposition 1. *Etant donnés* $\binom{n}{p}$ *nombres quelconques*

$$(2) \qquad A_{i_1 \ldots i_p}$$

numérotés par $i_1 < \ldots < i_p$, *on leur associe une fonctionnelle antisymétrique* A *unique dont ces nombres sont les coefficients stricts.*

Démonstration. L'unicité vient d'être établie. On démontre donc l'existence. On définit par (1) n^p nombres $A_{i_1 \ldots i_p}$ pour tous les $i_1, \ldots, i_p$, et on considère la fonctionnelle multilinéaire

$$(3) \qquad A = A_{i_1 \ldots i_p} \mathbf{e}^{i_1} \otimes \ldots \otimes \mathbf{e}^{i_p}.$$

Si A ainsi définie est antisymétrique, ses coefficients stricts sont évidemment les nombres (2). Aussi il faut établir la propriété d'antisymétrie de (3).

Aux termes de la proposition 1 de la leçon précédente, une condition suffisante en est la validité des relations

$$(4) \qquad A_{i_{\sigma(1)} \ldots i_{\sigma(p)}} = \varepsilon_\sigma A_{i_1 \ldots i_p}$$

pour les coefficients $A_{i_1 \ldots i_p}$ de la fonctionnelle, σ étant une permutation quelconque de degré p. On suppose, sans restreindre la généralité, que les indices $i_1, \ldots, i_p$ sont distincts deux à deux (sinon les deux membres de (4) sont nuls).

S'il en est ainsi, on a par définition

$$A_{i_1 \ldots i_p} = \varepsilon_\tau A_{i_{\tau(1)} \ldots i_{\tau(p)}},$$

avec τ une permutation telle que $i_{\tau(1)} < \ldots < i_{\tau(p)}$. De même,

$$A_{i_{\sigma(1)} \ldots i_{\sigma(p)}} = \varepsilon_\rho A_{i_{\rho(\sigma(1))} \ldots i_{\rho(\sigma(p))}},$$

avec ρ une permutation telle que $i_{\rho(\sigma(1))} < \ldots < i_{\rho(\sigma(p))}$. Or les nombres $i_{\tau(1)}, \ldots, i_{\tau(p)}$ et $i_{\rho(\sigma(1))}, \ldots, i_{\rho(\sigma(p))}$ représentant les indices $i_1, \ldots, i_p$ rangés dans l'ordre croissant sont identiques. Par conséquent,

$$\tau(1) = \rho(\sigma(1)), \ldots, \tau(p) = \rho(\sigma(p)),$$

i.e. $\tau = \rho\sigma$. Aussi $\varepsilon_\tau = \varepsilon_\rho \varepsilon_\sigma$ et, partant,

$$A_{i_1 \ldots i_p} = \varepsilon_\sigma A_{i_{\sigma(1)} \ldots i_{\sigma(p)}}. \quad \square$$

Théorème 1. *Les produits extérieurs*

$$\mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_p}, \quad 1 \leqslant i_1 < \ldots < i_p \leqslant n, \tag{5}$$

constituent une base de l'espace $\wedge_p(\mathscr{V})$.

Démonstration. La proposition 8 de la leçon précédente fait qu'il suffit de démontrer l'indépendance linéaire des fonctionnelles (5).

Soit

$$\sum_{i_1 < \ldots < i_p} A_{i_1 \ldots i_p} \mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_p} = 0,$$

où $A_{i_1 \ldots i_p}$, $i_1 < \ldots < i_p$, sont des nombres. Il existe par la proposition 1 une fonctionnelle antisymétrique A à coefficients stricts $A_{i_1 \ldots i_p}$. Selon la proposition 8 mentionnée, A est définie par la formule

$$A = \frac{1}{p!} \sum_{i_1 < \ldots < i_p} A_{i_1 \ldots i_p} \mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_p},$$

si bien qu'elle est nulle par hypothèse. Dans ce cas,

$$A_{i_1 \ldots i_p} = A(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_p}) = 0.$$

Donc, les fonctionnelles (5) sont linéairement indépendantes. $\square$

Conséquence 1. *Toute fonctionnelle antisymétrique A est représentable de façon unique par*

$$A = \frac{1}{p!} \sum_{i_1 < \ldots < i_p} A_{i_1 \ldots i_p} \mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_p}. \quad \square$$

Conséquence 2.

$$\dim \Lambda_p(\mathscr{V}) = \binom{n}{p}. \quad \square \tag{6}$$

En particulier, on retrouve l'égalité

$$\Lambda_p(\mathscr{V}) = 0, \quad p > n.$$

* * *

On passe de la base $\mathbf{e}^1, \ldots, \mathbf{e}^n$ à $\mathbf{e}^{1'}, \ldots, \mathbf{e}^{n'}$. Sous l'hypothèse ordinaire

$$\mathbf{e}^{i'} = c_i^{i'} \mathbf{e}^i,$$

on a

$$\mathbf{e}^{i'_1} \otimes \ldots \otimes \mathbf{e}^{i'_p} = c_{i_1}^{i'_1} \ldots c_{i_p}^{i'_p} \mathbf{e}^{i_1} \otimes \ldots \otimes \mathbf{e}^{i_p}.$$

Aussi

$$\mathbf{e}^{i'_1} \wedge \ldots \wedge \mathbf{e}^{i'_p} = \operatorname{Alt}(\mathbf{e}^{i'_1} \otimes \ldots \otimes \mathbf{e}^{i'_p}) = c_{[i_1}^{i'_1} \ldots c_{i_p]}^{i'_p} \mathbf{e}^{i_1} \otimes \ldots \otimes \mathbf{e}^{i_p}$$

(leçon 7, prop. 4), donc

$$\mathbf{e}^{i'_1} \wedge \ldots \wedge \mathbf{e}^{i'_p} = p! \sum_{i_1 < \ldots < i_p} c_{[i_1}^{i'_1} \ldots c_{i_p]}^{i'_p} \mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_p}.$$

Le nombre $p!c_{[i_1}^{i'_1} \ldots c_{i_p]}^{i'_p}$ est égal (voir leçon 7, lemme 1) au mineur

$$C\begin{pmatrix} i'_1 & \ldots & i'_p \\ i_1 & \ldots & i_p \end{pmatrix}$$

de la matrice de passage $C = (c_i^{i'})$ qui est constitué d'éléments en colonnes $i_1 < \ldots < i_p$ et en lignes $i'_1 < \ldots < i'_p$. Aussi on récrit finalement la formule de transformation des bases de l'espace $\Lambda_p(\mathscr{V})$:

$$\mathbf{e}^{i'_1} \wedge \ldots \wedge \mathbf{e}^{i'_p} = \sum_{i_1 < \ldots < i_p} C\begin{pmatrix} i'_1 & \ldots & i'_p \\ i_1 & \ldots & i_p \end{pmatrix} \mathbf{e}^{i_1} \wedge \ldots \wedge \mathbf{e}^{i_p}, \tag{7}$$

où $i'_1 < \ldots < i'_p$.

* * *

Il est naturel qu'on transpose tous les résultats obtenus aux tenseurs de type $(0, p)$, i.e. aux fonctionnelles multilinéaires $A : \xi^1, \ldots, \xi^p \mapsto A(\xi^1, \ldots, \xi^p)$ de covecteurs de degré p. La seule différence est la position des indices: les indices inférieurs sont élevés, et vice versa. En particulier, les fonctionnelles A de type $(0, p)$ ont pour coefficients $A^{i_1 \ldots i_p}$, la base de l'espace $\Lambda^p(\mathscr{V}) = \Lambda_p(\mathscr{V}')$

des fonctionnelles antisymétriques de type $(0, p)$ est formée de produits extérieurs

$$\mathbf{e}_{i_1} \wedge \ldots \wedge \mathbf{e}_{i_p} = \mathrm{Alt}\,(\mathbf{e}_{i_1} \otimes \ldots \otimes \mathbf{e}_{i_p}), \quad \text{où} \quad i_1 < \ldots < i_p,$$

et toute fonctionnelle antisymétrique admet dans cette base la décomposition

$$A = p! \sum_{i_1 < \ldots < i_p} A^{i_1 \ldots i_p} \mathbf{e}_{i_1} \wedge \ldots \wedge \mathbf{e}_{i_p}.$$

Définition 2. Les produits extérieurs

$$\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p, \ \mathbf{x}_1, \ldots, \mathbf{x}_p \in \mathscr{V},$$

de p vecteurs s'appellent les *multivecteurs de degré* p ou les *p-vecteurs*.

Si $p = 0$, les multivecteurs sont des nombres ($\in \mathbb{K}$), et si $p = 1$, ce sont des vecteurs. On désignera par $\mathbf{A}^p(\mathscr{V})$ l'ensemble des p-vecteurs de l'espace $\mathscr{V}$. $\mathbf{A}^p(\mathscr{V})$ n'est en général pas un espace vectoriel (la somme de deux p-vecteurs n'étant nécessairement pas un p-vecteur).

Les produits extérieurs $\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$ satisfont aux mêmes formules que les produits extérieurs $\xi^1 \wedge \ldots \wedge \xi^p$ de covecteurs (qu'on pourrait appeler les « multicovecteurs »). En particulier

$$\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p = \mathbf{x}_{[1} \otimes \ldots \otimes \mathbf{x}_{p]} \tag{8}$$

et

$$\begin{aligned}(\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p)(\xi^1, \ldots, \xi^p) &= \xi^1(\mathbf{x}_{[1}) \ldots \xi^p(\mathbf{x}_{p]}) = \\ &= \xi^{[1}(\mathbf{x}_1) \ldots \xi^{p]}(\mathbf{x}_p) = \begin{vmatrix} \xi^1(\mathbf{x}_1) & \ldots & \xi^p(\mathbf{x}_1) \\ \cdot & \cdot & \cdot \\ \xi^1(\mathbf{x}_p) & \ldots & \xi^p(\mathbf{x}_p) \end{vmatrix}\end{aligned} \tag{9}$$

pour $\xi^1, \ldots, \xi^p$ quelconques. Cela étant, *l'égalité*

$$\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p = 0$$

a lieu si et seulement si $\mathbf{x}_1, \ldots, \mathbf{x}_p$ *sont linéairement dépendants.* (Cf. leçon 8, prop. 7.) □

On ferait bien de noter qu'on a pour n vecteurs quelconques

$$\begin{aligned} \mathbf{x}_1 &= x_1^1 \mathbf{e}_1 + \ldots + x_1^n \mathbf{e}_n, \\ &\cdots\cdots\cdots\cdots \\ \mathbf{x}_n &= x_n^1 \mathbf{e}_1 + \ldots + x_n^n \mathbf{e}_n \end{aligned}$$

l'égalité

$$\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_n = \begin{vmatrix} x_1^1 & \ldots & x_1^n \\ \cdot & \cdot & \cdot \\ x_n^1 & \ldots & x_n^n \end{vmatrix} (\mathbf{e}_1 \wedge \ldots \wedge \mathbf{e}_n). \tag{10}$$

En effet, s'il y a dépendance linéaire des $\mathbf{x}_1, \ldots, \mathbf{x}_n$, l'égalité est évidente (on a à droite et à gauche les n-vecteurs nuls). Mais si $\mathbf{x}_1, \ldots, \mathbf{x}_n$ sont indépendants (et constituent donc une base de $\mathscr{V}$), seules les notations distinguent (10) d'avec l'analogue « vectoriel » pour $p = n$ de la formule (7).

Lorsque $n = 3$, la formule (10) est identique aux notations près à la formule (3) de I.13. Aussi on s'attend naturellement à ce que les p-vecteurs au sens de la définition 2 ne soient en fait autres que les p-vecteurs introduits dans I.12 (i.e. les classes de familles équivalentes de vecteurs), i.e. à ce que l'égalité

$$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_p = \mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_p$$

ait lieu si et seulement si les familles $\mathbf{x}_1, \ldots, \mathbf{x}_p$ et $\mathbf{y}_1, \ldots, \mathbf{y}_p$ sont unimodulairement équivalentes (cf. I.12, prop. 2). Il se trouve qu'il en est bien ainsi. S'agissant des familles liées, la propriété découle immédiatement des résultats ci-dessus, si bien qu'on se borne, sans nuire à la généralité, aux familles libres de vecteurs.

Théorème 2. *Etant données deux familles libres de vecteurs* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *et* $\mathbf{y}_1, \ldots, \mathbf{y}_p$, *l'égalité*

$$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_p = \mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_p$$

a lieu si et seulement si les familles sont unimodulairement équivalentes, i.e. si

$$\begin{aligned} \mathbf{y}_1 &= c_1^1 \mathbf{x}_1 + \cdots + c_1^p \mathbf{x}_p, \\ &\cdots\cdots\cdots\cdots \\ \mathbf{y}_p &= c_p^1 \mathbf{x}_1 + \cdots + c_p^p \mathbf{x}_p, \end{aligned} \tag{11}$$

où

$$\begin{vmatrix} c_1^1 & \ldots & c_1^p \\ \cdot & \cdot\cdot\cdot & \cdot \\ c_p^1 & \ldots & c_p^p \end{vmatrix} = 1. \tag{12}$$

L'affirmation dans un sens découle immédiatement de la formule (10). En effet, si l'on a (11), alors les deux familles sont des bases d'un même sous-espace de dimension p. Aussi, on a par (10)

$$\mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_p = \Delta\,(\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_p), \tag{13}$$

Δ étant le déterminant (12). Il reste à observer que $\Delta = 1$ par hypothèse. □

L'affirmation dans l'autre sens est sensiblement plus difficile à démontrer. On fait quelques remarques préliminaires.

* * *

Comme pour les fonctionnelles multilinéaires arbitraires, on définit pour les fonctionnelles antisymétriques les notions de rang et d'espace associé. On conçoit que s'agissant des fonctionnelles

antisymétriques, il est beaucoup plus naturel que ces notions utilisent le produit extérieur au lieu du produit tensoriel.

Soit A une fonctionnelle antisymétrique quelconque de degré p (on suppose, pour fixer les idées, que c'est une fonctionnelle de covecteurs).

Définition 3. On dit que la fonctionnelle A *s'exprime de façon extérieure* au moyen des vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_r$ si elle est combinaison linéaire des produits extérieurs $\mathbf{x}_{i_1} \wedge \ldots \wedge \mathbf{x}_{i_p}$, où $1 \leqslant i_1, \ldots, i_p \leqslant r$. Cf. définition 4 de la leçon 6.

On appelle *rang extérieur* de la fonctionnelle antisymétrique A le nombre r tel que

1) il existe une famille de r vecteurs dont A est combinaison extérieure ;

2) si A s'exprime de manière extérieure par une famille de vecteurs, la famille renferme au moins r éléments.

Chose remarquable, ces définitions sont en fait superflues car *la fonctionnelle antisymétrique A s'exprime de façon extérieure en fonction des vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_r$ si et seulement si elle en est combinaison tensorielle* (si bien que *le rang extérieur de A est en fait identique à son rang*). En effet, si

$$A = a^{i_1 \ldots i_p} \mathbf{x}_{i_1} \otimes \ldots \otimes \mathbf{x}_{i_p},$$

on obtient par application alternée

$$A = a^{i_1 \ldots i_p} \mathbf{x}_{i_1} \wedge \ldots \wedge \mathbf{x}_{i_p}. \tag{14}$$

Inversement, si l'on a (14), alors

$$A = a^{i_1 \ldots i_p} \mathbf{x}_{[i_1} \otimes \ldots \otimes \mathbf{x}_{i_p]}$$

selon (8). □

Mais la notion de rang extérieur n'est pas inutile. On conçoit en effet que le rang extérieur d'une fonctionnelle antisymétrique non nulle est au moins égal à son degré (sinon les facteurs de chaque terme de (14) seraient non distincts deux à deux). On a donc de même pour le rang d'une fonctionnelle antisymétrique non nulle la

Proposition 2. *Le rang r d'une fonctionnelle antisymétrique $A \in \in \wedge^p(\mathscr{V})$ non nulle est au moins égal à son degré :*

$$p \leqslant r. \ \square$$

Dans la suite, nous utiliserons plus d'une fois ce résultat important.

* * *

La première application de la proposition 2 sera la caractérisation des multivecteurs dans la classe de toutes les fonctionnelles antisymétriques, à savoir la

Proposition 3. *Une fonctionnelle antisymétrique* $A \in \wedge^p(\mathscr{V}')$ *est un multivecteur si et seulement si son rang* r *est égal à son degré*:

$$p = r.$$

Démonstration. Si $A = \mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$, alors $r \leqslant p$ de façon évidente. D'après la proposition 2, on a donc $r = p$.

Inversement, soit $r = p$. La fonctionnelle A s'écrit

$$A = a^{i_1 \ldots i_p} \mathbf{x}_{i_1} \wedge \ldots \wedge \mathbf{x}_{i_p},$$

$\mathbf{x}_1, \ldots, \mathbf{x}_p$ étant une base de l'espace associé. On a montré plus d'une fois que A s'exprime dans ce cas par la formule

$$A = a^{[1 \ldots p]} (\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p),$$

donc par

$$A = \mathbf{y}_1 \wedge \ldots \wedge \mathbf{y}_p,$$

où $\mathbf{y}_1 = a^{[1 \ldots p]} \mathbf{x}_1$, $\mathbf{y}_2 = \mathbf{x}_2, \ldots, \mathbf{y}_p = \mathbf{x}_p$. □

Conséquence. *L'espace associé à un multivecteur* $\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p \neq 0$ *est l'enveloppe linéaire* $[\mathbf{x}_1, \ldots, \mathbf{x}_p]$ *des vecteurs* $\mathbf{x}_1, \ldots, \mathbf{x}_p$.

* * *

Fort de ces résultats, on établit les conditions d'égalité de deux multivecteurs.

Proposition 4. *Etant donnés deux p-vecteurs non nuls*

$$\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p \quad \textit{et} \quad \mathbf{y}_1 \wedge \ldots \wedge \mathbf{y}_p,$$

quatre affirmations suivantes sont équivalentes:

a) *les p-vecteurs donnés sont proportionnels, i.e. il existe un nombre* $k \neq 0$ *tel que*

$$\mathbf{y}_1 \wedge \ldots \wedge \mathbf{y}_p = k (\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p);$$

b) *les vecteurs* $\mathbf{x}_1, \ldots, \mathbf{x}_p$ *et* $\mathbf{y}_1, \ldots, \mathbf{y}_p$ *ont même enveloppe linéaire*:

$$[\mathbf{y}_1, \ldots, \mathbf{y}_p] = [\mathbf{x}_1, \ldots, \mathbf{x}_p];$$

c) *les familles* $\mathbf{x}_1, \ldots, \mathbf{x}_p$ *et* $\mathbf{y}_1, \ldots, \mathbf{y}_p$ *sont linéairement équivalentes, i.e. on a les égalités*

$$\begin{aligned} \mathbf{y}_1 &= c_1^1 \mathbf{x}_1 + \ldots + c_1^p \mathbf{x}_p, \\ &\ldots\ldots\ldots\ldots \\ \mathbf{y}_p &= c_p^1 \mathbf{x}_1 + \ldots + c_p^p \mathbf{x}_p, \end{aligned}$$

où

$$\begin{vmatrix} c_1^1 & \ldots & c_1^p \\ \cdot & \cdot & \cdot \\ c_p^1 & \ldots & c_p^p \end{vmatrix} \neq 0;$$

d) *les espaces associés aux p-vecteurs* $\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$ *et* $\mathbf{y}_1 \wedge \ldots \wedge \mathbf{y}_p$ *sont identiques.*

Démonstration. Soit

$$\mathbf{y}_1 \wedge \ldots \wedge \mathbf{y}_p = k\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p.$$

Les fonctionnelles $\mathbf{y}_1 \wedge \ldots \wedge \mathbf{y}_p$ et $k\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$ étant égales, leurs espaces associés coïncident. Aussi, la conséquence ci-dessus implique

$$[\mathbf{y}_1, \ldots, \mathbf{y}_p] = [k\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p].$$

Mais il est clair que

$$[k\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p] = [\mathbf{x}_1, \ldots, \mathbf{x}_p],$$

si bien que la même conséquence entraîne l'égalité des espaces associés aux fonctionnelles $k\mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \ldots \wedge \mathbf{x}_p$ et $\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$. Ainsi, a) $\Rightarrow$ b) et a) $\Rightarrow$ d).

L'équivalence de b) et c) est évidente (nous en avons parlé dans la leçon première).

Si l'affirmation c) est vérifiée, alors les multivecteurs $\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$ et $\mathbf{y}_1 \wedge \ldots \wedge \mathbf{y}_p$ sont reliés par (13), i.e. on a a).

La condition d) implique c) car l'espace associé à la fonctionnelle $\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$ a pour base les vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_p$.

Par conséquent, il y a équivalence des conditions a), b), c) et d.) □

Conséquence. *On a, entre les classes de p-vecteurs non nuls proportionnels et les sous-espaces p-dimensionnels de l'espace* $\mathcal{V}$, *une correspondance bijective canonique qui associe à chaque sous-espace* $\mathcal{P}$ *le produit extérieur* $\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$ *de vecteurs de sa base* $\mathbf{x}_1, \ldots, \mathbf{x}_p$ *quelconque et à chaque p-vecteur* $\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$ *le sous-espace* $[\mathbf{x}_1, \ldots, \mathbf{x}_p]$.

La **démonstration du théorème 2** ne présente maintenant aucune difficulté. Il est connu que si l'on a (11) et (12) à la fois, alors $\mathbf{y}_1 \wedge \ldots \wedge \mathbf{y}_p = \mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$. Inversement, si $\mathbf{y}_1 \wedge \ldots \wedge \mathbf{y}_p = \mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$, alors on a (11) par suite de la proposition 4 et, partant, l'égalité (13), où $\Delta = 1$. □

LEÇON 10

Théorème de Cartan. — Relations de Plücker. — Coordonnées plückériennes des sous-espaces. — Plans dans l'espace affine. — Plans dans l'espace projectif et leurs coordonnées.

La proposition 3 de la leçon précédente a fourni une condition nécessaire et suffisante pour qu'une fonctionnelle antisymétrique soit un multivecteur. Ce critère est peu efficace dans la pratique. Il en existe un autre, plus commode. On démontre au préalable le *théorème d'Elie Cartan* (relatif à la divisibilité d'une fonctionnelle antisymétrique par un multivecteur):

Proposition 1. *Soit* $\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_r \neq 0$ *et une fonctionnelle antisymétrique* A *de degré* $p \geqslant r$. *Il existe une fonctionnelle antisymétrique* B *de degré* $p - r$ *telle que*

$$A = B \wedge \mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_r$$

si et seulement si

$$A \wedge \mathbf{x}_1 = 0, \ldots, A \wedge \mathbf{x}_r = 0. \tag{1}$$

Démonstration. Si $A = B \wedge \mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_r$, alors le produit extérieur

$$A \wedge \mathbf{x}_s = B \wedge \mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_r \wedge \mathbf{x}_s$$

contient pour tout $s = 1, \ldots, r$ deux facteurs identiques $\mathbf{x}_s$ et vaut donc 0.

Inversement, soit (1). Les vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_r$ étant linéairement indépendants par hypothèse sont complétés en une base

$$\mathbf{e}_1 = \mathbf{x}_1, \ldots, \mathbf{e}_r = \mathbf{x}_r, \mathbf{e}_{r+1}, \ldots, \mathbf{e}_n$$

de l'espace $\mathcal{V}$. Soit

$$A = p! \sum_{1 \leqslant i_1 < \ldots < i_p \leqslant n} A^{i_1 \ldots i_p} \mathbf{e}_{i_1} \wedge \ldots \wedge \mathbf{e}_{i_p} \tag{2}$$

la décomposition de A dans la base associée de l'espace $\wedge_p(\mathcal{V})$ (voir théorème 1, leçon 9). On a pour tout $s = 1, \ldots, r$

$$A \wedge \mathbf{x}_s = p! \sum_{1 \leqslant i_1 < \ldots < i_p \leqslant n} A^{i_1 \ldots i_p} \mathbf{e}_{i_1} \wedge \ldots \wedge \mathbf{e}_{i_p} \wedge \mathbf{e}_s.$$

Si s est égal à l'un des nombres $i_1, \ldots, i_p$, alors $\mathbf{e}_{i_1} \wedge \ldots \wedge \mathbf{e}_{i_p} \wedge \mathbf{e}_s = 0$. On se limite donc dans la somme pour $A \wedge \mathbf{x}_s$ aux termes dont tous les indices $i_1, \ldots, i_p \neq s$:

$$(2) \qquad A \wedge \mathbf{x}_s = p! \sum_{\substack{i_1 < \ldots < i_p \\ i_1, \ldots, i_p \neq s}} A^{i_1 \ldots i_p} \mathbf{e}_{i_1} \wedge \ldots \wedge \mathbf{e}_{i_p} \wedge \mathbf{e}_s.$$

Mais on sait que si $i_1, \ldots, i_p \neq s$, tous les multivecteurs de la forme

$$\mathbf{e}_{i_1} \wedge \ldots \wedge \mathbf{e}_{i_p} \wedge \mathbf{e}_s, \quad 1 \leqslant i_1 < \ldots < i_p \leqslant n,$$

sont linéairement indépendants. Aussi $A \wedge \mathbf{x}_s = 0$ entraîne la nullité de tous les coefficients $A^{i_1 \ldots i_p}$ de (3).

On vient de démontrer qu'avec les conditions (1), seuls sont non nuls les coefficients $A^{i_1 \ldots i_p}$ de la décomposition (2) pour lesquels chaque $i_1 < \ldots < i_p$ est égal à un $s = 1, \ldots, r$ et un seul, i.e. pour lesquels $i_1 = 1, \ldots, i_r = r$. Aussi

$$A = B \wedge \mathbf{e}_1 \wedge \ldots \wedge \mathbf{e}_r,$$

où

$$B = (-1)^{r(p-r)} p! \sum_{r < i_{r+1} < \ldots < i_p} A^{1 \ldots r i_{r+1} \ldots i_p} \mathbf{e}_{i_{r+1}} \wedge \ldots \wedge \mathbf{e}_{i_p}. \ \square$$

Proposition 2. *Une fonctionnelle antisymétrique* $A \in \wedge^p(\mathscr{V})$ *est un multivecteur si et seulement si*

$$A \wedge \mathbf{x} = 0$$

pour tout vecteur $\mathbf{x}$ *associé.*

Démonstration. Si $A = \mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$, alors $\mathbf{x}$ s'exprime linéairement en fonction des vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_p$ (car $\mathbf{x}_1, \ldots, \mathbf{x}_p$ engendrent l'espace associé), donc

$$A \wedge \mathbf{x} = \mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p \wedge \mathbf{x} = 0.$$

Inversement, si $A \wedge \mathbf{x} = 0$ pour tout $\mathbf{x}$ associé, alors $A \wedge \mathbf{x} = 0$ pour tout $\mathbf{x}$ de l'espace associé. En particulier, on a

$$A \wedge \mathbf{x}_1 = 0, \ \ldots, \ A \wedge \mathbf{x}_r = 0$$

si $\mathbf{x}_1, \ldots, \mathbf{x}_r$ est une base de l'espace associé. Aux termes de la proposition 1, il existe donc une fonctionnelle B de degré $p - r$ telle que

$$A = B \wedge \mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_r.$$

La chose n'est certes possible que pour $p \geqslant r$. Puisqu'on a toujours $p \leqslant r$ (leçon 9, prop. 2), cela démontre que $p = r$, i.e. que la fonctionnelle A est un multivecteur (leçon 9, prop. 3). $\square$

* * *

La propriété d'antisymétrie fait que tout vecteur associé à la fonctionnelle A est défini (en tant qu'élément de $(\mathscr{T}^{\cdot\prime})'$) par la formule

$$\mathbf{x}: \xi \longmapsto A\ (\xi,\ \beta^2,\ \ldots,\ \beta^p),$$

avec $\beta^2, \ldots, \beta^p$ certains covecteurs fixes, et a donc pour coordonnées

$$x^i = A^{ij_2\ldots j_p}\beta^2_{j_2}\ \ldots\ \beta^p_{j_p}.$$

C'est pourquoi les coordonnées (les coefficients) de la fonctionnelle $A \wedge \mathbf{x} = \mathrm{Alt}\ (A \otimes \mathbf{x})$ sont données par

$$(A \wedge \mathbf{x})^{i_1 i_2 \ldots i_{p+1}} = A^{[i_1 \ldots i_p} x^{i_{p+1}]} = A^{[i_1 \ldots i_p} A^{i_{p+1}]\, j_2 \ldots j_p} \beta^2_{j_2}\ \ldots\ \beta^p_{j_p}.$$

Ces expressions sont nulles pour tous les $\beta^2_{j_2}, \ldots, \beta^p_{j_p}$ si et seulement si

$$A^{[i_1 \ldots i_p} A^{i_{p+1}]\, j_2 \ldots j_p} = 0$$

pour tous les $i_1, \ldots, i_p, i_{p+1}, j_2, \ldots, j_p$ (si, par exemple, $A^{[i_1 \ldots i_p} A^{i_{p+1}] j_2^0 \ldots j_p^0} \neq 0$, alors $(A \wedge \mathbf{x})^{i_1 \ldots i_p l_{p+1}} \neq 0$ pour $\beta^2_{j_2} = \delta^{j_2^0}_{j_2}, \ldots, \beta^p_{j_p} = \delta^{j_p^0}_{j_p}$). En raison de symétrie on désigne i_{p+1} par j_1 et on aboutit à la

Proposition 3. *Une fonctionnelle antisymétrique $A \in \bigwedge^p(\mathscr{T}^{\cdot})$ est un multivecteur si et seulement si ses coefficients vérifient les relations*

$$A^{[i_1 \ldots i_p} A^{j_1]\, j_2 \ldots j_p} = 0 \tag{4}$$

pour tous les $i_1, \ldots, i_p, j_1, j_2, \ldots, j_p$. □

Les égalités (4) sont connues sous le nom de *relations de Plücker*.

Exemple. Les relations (4) s'écrivent pour $p = 2$

$$A^{[i_1 i_2} A^{j_1]\, j_2} = 0,$$

i.e.

$$A^{i_1 i_2} A^{j_1 j_2} + A^{i_2 j_1} A^{i_1 j_2} + A^{j_1 i_1} A^{i_2 j_2} - A^{i_2 i_1} A^{j_1 j_2} - A^{j_1 i_2} A^{i_1 j_2} - A^{i_1 j_1} A^{i_2 j_2} = 0.$$

Etant donnée l'antisymétrie, le premier terme est égal au quatrième, le deuxième au cinquième et le troisième au sixième, si bien qu'on a par réduction des termes semblables et division par 2:

$$A^{i_1 i_2} A^{j_1 j_2} + A^{i_2 j_1} A^{i_1 j_2} + A^{j_1 i_1} A^{i_2 j_2} = 0. \tag{5}$$

Si $i_1 = i_2$, le premier terme est nul et les deux termes restants ne diffèrent que par le signe. Dans ce cas, les relations (5) sont vérifiées automatiquement (en vertu de l'antisymétrie). Le cas des indices i_1, i_2, j_1, j_2 égaux deux à deux se traite de façon analogue. Seules les relations (5) à i_1, i_2, j_1, j_2 distincts sont donc non triviales.

La chose est impossible pour $n = 3$, si bien que toutes les relations de Plücker sont triviales pour $n = 3$ (et $p = 2$), i.e. *dans l'espace de dimension* 3 *toute fonctionnelle antisymétrique est un bivecteur*. C'est la raison pour laquelle nous avons réussi dans le Livre I

à transformer pour $n = 3$ l'ensemble de bivecteurs en un espace vectoriel.

Quand $n = 4$, il existe une seule relation de Plücker non triviale, à savoir

$$A^{12}A^{34} + A^{23}A^{14} + A^{13}A^{42} = 0.$$

Aussi les bivecteurs ne forment pas pour $n = 4$ un espace vectoriel (la somme de deux bivecteurs n'est en général pas un bivecteur).

On montre de manière analogue pour tout n que toutes les relations de Plücker sont triviales pour $p = n - 1$, i.e. que *toute fonctionnelle antisymétrique de degré $n - 1$ est un $n - 1$-vecteur*. Les coefficients stricts $A^{i_1 \ldots i_{n-1}}$, $1 \leqslant i_1 < \ldots < i_{n-1} \leqslant n$, de la fonctionnelle s'écrivent évidemment $A^{1 \ldots \hat{i} \ldots n}$, le signe $\hat{}$ particularisant l'indice à omettre. La notation commode est en l'occurrence $(-1)^i B_i$.

Remarque. Puisque les nombres B_i s'interprètent comme coordonnées d'un covecteur, il y a une bijection entre les $n - 1$-vecteurs et les covecteurs (i.e. les 1-vecteurs). Il se trouve qu'une correspondance analogue relie pour tout p les p-vecteurs et les $n - p$-covecteurs. Cette correspondance dépend en général de la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$, mais cette dépendance est assez faible. Elle est notamment la même pour toutes les bases unimodulairement équivalentes, i.e. pour les bases définissant un même n-vecteur:

$$E_0 = \mathbf{e}_1 \wedge \ldots \wedge \mathbf{e}_n.$$

Cette bijection associe à chaque $n - p$-covecteur $B = \beta^1 \wedge \ldots \wedge \beta^{n-p}$ un p-vecteur A défini en tant que fonctionnelle antisymétrique par la formule

$$A(\xi^1, \ldots, \xi^p) = E_0(\beta^1, \ldots, \beta^{n-p}, \xi^1, \ldots, \xi^p).$$

Quel que soit le n-vecteur E_0, elle est définie à un facteur de proportionnalité près, i.e. elle l'est comme correspondance entre les classes de p-vecteurs proportionnels et les classes de $n - p$-covecteurs proportionnels. Par identification de ces classes et des sous-espaces de $\mathscr{V}$ et $\mathscr{V}'$ (voir conséquence de la proposition 4, leçon 9) on établit une correspondance entre chaque sous-espace $\mathscr{P} \subset \mathscr{V}$ et son annulateur $\mathscr{P}^\circ$.

Nous montrerons en temps et lieu que dans l'espace euclidien on identifie les covecteurs et les vecteurs, donc les $n - p$-covecteurs et les $n - p$-vecteurs. Il en résulte que dans l'espace euclidien orienté on définit une bijection entre les p-vecteurs et les $n - p$-vecteurs. Lorsque $n = 3$ et $p = 2$, cette bijection coïncide avec la correspondance de la définition 4 de I.15.

Malheureusement, le temps nous manque pour entrer dans le détail de ce problème intéressant.

On conçoit que les relations de Plücker sont triviales pour $p = n$. D'ailleurs, cela résulte immédiatement du théorème 1 de la leçon 9 (ou plutôt de sa variante vectorielle) qui veut que l'espace $\mathsf{A}^n(\mathscr{V})$ soit de dimension 1 et engendré par tout n-vecteur $\mathbf{e}_1 \wedge \ldots \wedge \mathbf{e}_n \neq \neq 0$. Aussi chaque fonctionnelle antisymétrique de degré n est un n-vecteur de la forme $a\mathbf{e}_1 \wedge \ldots \wedge \mathbf{e}_n$.

Ainsi,

$$\mathsf{A}^{n-1}(\mathscr{V}) = \Lambda^{n-1}(\mathscr{V}) \quad \text{et} \quad \mathsf{A}^n(\mathscr{V}) = \Lambda^n(\mathscr{V}).$$

On a, de plus,

$$\mathsf{A}^0(\mathscr{V}) = \Lambda^0(\mathscr{V}) = \mathbb{K} \quad \text{et} \quad \mathsf{A}^1(\mathscr{V}) = \Lambda^1(\mathscr{V}) = \mathscr{V}.$$

Le temps est venu d'appliquer les résultats obtenus à la géométrie de l'espace.

* * *

Définition 1. Un p-vecteur A non nul s'appelle *p-vecteur directeur* d'un sous-espace $\mathscr{P} \subset \mathscr{V}$ p-dimensionnel si $\mathscr{P}$ est le sous-espace associé à A. On dit également des vecteurs $\mathbf{x} \in \mathscr{P}$ qu'ils sont *parallèles* au p-vecteur A et on note $\mathbf{x} \parallel A$.

Etant donnée la conséquence de la proposition 4 de la leçon 9, chaque p-vecteur directeur est le produit extérieur $\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$ de vecteurs d'une base de $\mathscr{P}$, si bien qu'il est bien défini, à un facteur de proportionnalité près, par $\mathscr{P}$ (et définit certes $\mathscr{P}$ de façon unique).

Définition 2. Les coordonnées $A^{i_1 \ldots i_p}$ d'un p-vecteur directeur quelconque de $\mathscr{P} \subset \mathscr{V}$ de dimension p s'appellent les *coordonnées plückériennes* de $\mathscr{P}$. Elles sont définies (pour une base fixe de $\mathscr{V}$) à un facteur de proportionnalité près, i.e. ce sont les coordonnées homogènes.

Dans une base $\mathbf{x}_1, \ldots, \mathbf{x}_p$ quelconque de $\mathscr{P}$, les coordonnées plückériennes de $\mathscr{P}$ sont représentées par les formules

$$\rho A^{i_1 \ldots i_p} = x_1^{[i_1} \ldots x_p^{i_p]}$$

ou, ce qui revient au même, par

$$\rho A^{i_1 \ldots i_p} = \begin{vmatrix} x_1^{i_1} & \ldots & x_p^{i_p} \\ \cdot & \cdot & \cdot \\ x_p^{i_1} & \ldots & x_p^{i_p} \end{vmatrix},$$

ρ étant un facteur de proportionnalité arbitraire.

Un ensemble de nombres $A^{i_1 \ldots i_p}$ est un ensemble de coordonnées plückériennes d'un sous-espace $\mathscr{P}$ de dimension p si et seulement si

1) $A^{i_1 \ldots i_p}$ dépendent des indices de façon antisymétrique, i.e. toute permutation $\sigma \in S_p$ remplit l'égalité

$$A^{i_{\sigma(1)} \ldots i_{\sigma(p)}} = \varepsilon_\sigma A^{i_1 \ldots i_p};$$

2) quels que soient les indices $i_1, \ldots, i_p, j_1, \ldots, j_p$, on a la relation de Plücker

$$A^{[i_1 \ldots i_p} A^{j_1] j_2 \ldots j_p} = 0 .$$

Il s'agit d'une reformulation évidente des résultats connus.

* * *

Soient $\mathcal{A}$ un espace affine de dimension n (voir I.5) et $\mathcal{A}$ l'espace vectoriel attaché. On imite la définition de la droite et du plan (I.5 et I.6) et on introduit la

Définition 3. Soient M_0 un point quelconque de $\mathcal{A}$ et A un p-vecteur non nul arbitraire de $\mathbf{A}^p(\mathcal{V})$. On appelle le *plan de dimension p* parallèle à A par M_0 l'ensemble de tous les points $M \in \mathcal{A}$ tels que $\overrightarrow{M_0M} \parallel A$. On dit du p-vecteur A qu'il est un p-vecteur directeur du plan.

Si $p = 1$, le plan constitue une *droite* (voir I.5, déf. 7). C'est un *hyperplan* pour $p = n - 1$.

Lorsque $p = 0$, les plans sont des points de l'espace $\mathcal{A}$.

Si $A = \mathbf{a}_1 \wedge \ldots \wedge \mathbf{a}_p$ et si l'on fixe dans $\mathcal{A}$ un point O, la condition $\overrightarrow{M_0M} \parallel A$ équivaut à

$$\mathbf{x} = \mathbf{x}_0 + t^1\mathbf{a}_1 + \ldots + t^p\mathbf{a}_p, \tag{6}$$

avec $\mathbf{x}_0 = \overrightarrow{OM_0}$ et $\mathbf{x} = \overrightarrow{OM}$ les rayons vecteurs des points M_0 et M respectivement et $t^1, \ldots, t^p$ des nombres quelconques. L'égalité (6) constitue par définition l'*équation paramétrique vectorielle* du plan.

Dans un repère affine $O\mathbf{e}_1 \ldots \mathbf{e}_n$ quelconque, l'équation (6) est équivalente à n équations numériques

$$\begin{aligned} x^1 &= x_0^1 + t^1a_1^1 + \ldots + t^pa_p^1, \\ &\ldots\ldots\ldots\ldots\ldots \\ x^n &= x_0^n + t^1a_1^n + \ldots + t^pa_p^n \end{aligned} \tag{7}$$

appelées *équations paramétriques* (*en coordonnées*) du plan. On définit un plan en se donnant, au lieu d'un p-vecteur directeur A, le sous-espace $\mathcal{P} \subset \mathcal{V}$ correspondant dont les éléments sont parallèles à A (i.e. $\mathcal{P}$ en est l'espace associé). On dit des vecteurs de $\mathcal{P}$ qu'ils sont *parallèles* au plan considéré.

L'égalité (6) signifie qu'un point M de rayon vecteur $\mathbf{x}$ est dans le plan si et seulement si $\mathbf{x} - \mathbf{x}_0 \in \mathcal{P}$, i.e. si $\mathbf{x}$ appartient à la classe $\mathbf{x}_0 + \mathcal{P}$ de $\mathcal{V}$ suivant $\mathcal{P}$, ce qui légitime la

Définition 4. On dit qu'une partie $\mathcal{Q}$ de l'espace vectoriel $\mathcal{V}$ est une *variété linéaire* s'il existe un sous-espace $\mathcal{P} \subset \mathcal{V}$ (manifestement unique) tel que $\mathcal{Q}$ soit une classe de $\mathcal{V}$ suivant $\mathcal{P}$. La dimension de $\mathcal{P}$ est par définition la *dimension* de la variété linéaire $\mathcal{Q}$.

On dit donc que les plans de dimension p de l'espace affine $\mathcal{A}$ sont exactement les parties de $\mathcal{A}$ qui deviennent par la bijection

$$M \mapsto \mathbf{x} = \overrightarrow{OM}$$

les variétés linéaires de dimension p de l'espace $\mathcal{V}$. Autrement dit, le choix du point $O \in \mathcal{A}$ permet d'identifier les plans de l'espace affine $\mathcal{A}$ et les variétés linéaires de l'espace vectoriel $\mathcal{V}$.

Il est connu dès la leçon 4 qu'on définit les sous-espaces $\mathcal{P} \subset \mathcal{V}$ comme annulateurs des familles de covecteurs $\xi^1, \ldots, \xi^m \in \mathcal{V}'$, i.e. par les conditions

$$\xi^1(\mathbf{x}) = 0, \ldots, \xi^m(\mathbf{x}) = 0.$$

Il en résulte qu'un vecteur $\mathbf{x} \in \mathcal{V}$ appartient à la variété linéaire $\mathbf{x}_0 + \mathcal{P}$ si et seulement si

$$\xi^1(\mathbf{x}) = b^1, \ldots, \xi^m(\mathbf{x}) = b^m, \tag{8}$$

où $b^1 = \xi^1(\mathbf{x}_0), \ldots, b^m = \xi^m(\mathbf{x}_0)$. Cela signifie que les égalités (8) caractérisent les rayons vecteurs $\mathbf{x}$ des points du plan correspondant de l'espace $\mathcal{A}$.

En coordonnées, ces égalités s'écrivent

$$\begin{array}{c} \xi_1^1 x^1 + \ldots + \xi_n^1 x^n = b^1, \\ \cdots\cdots\cdots\cdots \\ \xi_1^m x^1 + \ldots + \xi_n^m x^n = b^m, \end{array} \tag{9}$$

i.e. c'est un système d'équations linéaires non homogènes.

Ainsi, on a le

Théorème 1. *Etant donné un système quelconque d'équations linéaires* (9), *les points de l'espace* $\mathcal{A}$ (*resp. les vecteurs de l'espace* $\mathcal{V}$) *qui satisfont en coordonnées* $x^1, \ldots, x^n$ *à* (9), *forment, quand ils existent, un plan* (*resp. une variété linéaire*), *tout plan* (*resp. toute variété linéaire*) *pouvant être obtenu de cette façon.* □

Les équations (9) constituent donc les *équations du plan* (resp. *de la variété linéaire*). La dimension du plan est $n - r$, r étant le rang de la matrice des coefficients du système (9) (voir leçon 4, th. 2).

Algébriquement, passer des équations (9) du plan aux équations paramétriques (7) c'est chercher la solution générale du système (9) (par le procédé décrit dans la leçon 2), et passer de (7) à (9) c'est écrire les équations dont la solution générale est de la forme (7).

Les vecteurs $\mathbf{a}_1, \ldots, \mathbf{a}_p$ étant linéairement indépendants par hypothèse, la matrice

$$\begin{pmatrix} a_1^1 & \ldots & a_p^1 \\ \cdot & \cdots & \cdot \\ a_1^n & \ldots & a_p^n \end{pmatrix}$$

de leurs coordonnées renferme un mineur non nul d'ordre p. Si l'on considère les égalités (7) correspondantes comme étant des équations en $t^1, \ldots, t^p$, on exprime par Cramer les inconnues $t^1, \ldots, t^p$ en fonction de $x^1, \ldots, x^n$. On substitue les valeurs obtenues dans $n - p$ équations restantes et on trouve pour $x^1, \ldots, x^p$ des équations de la forme (9) (avec $m = n - p$).

Ainsi, on constate une équivalence parfaite entre la théorie « géométrique » des plans dans l'espace affine n-dimensionnel et la théorie « algébrique » des systèmes d'équations linéaires non homogènes à n inconnues. Les deux théories disent la même chose, mais dans des langues différentes. Il faut savoir passer sans hésitation d'une langue à l'autre.

Exemple. L'unicité de la solution du système d'équations (9) signifie que le plan associé est de dimension 0 et constitue un point de l'espace $\mathscr{A}$. Le sous-espace $\mathscr{P} \subset \mathscr{V}$ correspondant est réduit au vecteur $\mathbf{0}$. Par conséquent, le système d'équations homogènes

$$\begin{aligned} &\xi_1^1 x^1 + \ldots + \xi_n^1 x^n = 0, \\ &\ldots\ldots\ldots\ldots \\ &\xi_1^m x^1 + \ldots + \xi_n^m x^n = 0 \end{aligned} \tag{10}$$

ne possède que la solution triviale (nulle) $(0, \ldots, 0)$. Inversement, si le système (10) admet seulement la solution triviale, le sous-espace $\mathscr{P}$ défini par (10) n'a pas d'autre élément à part le vecteur $\mathbf{0}$. Aussi chaque classe $\mathbf{x} + \mathscr{P}$ suivant $\mathscr{P}$ est formée du vecteur $\mathbf{x}$ seul et, partant, il y a unicité de la solution de (9). Ainsi, le système (9) (compatible) d'équations linéaires non homogènes a une solution unique si et seulement si le système (10) d'équations linéaires homogènes admet la solution triviale et cette solution seulement.

Géométriquement, si $\mathscr{P} = \mathbf{0}$, alors $\mathbf{x}_0 + \mathscr{P} = \mathbf{x}_0$ pour tout $\mathbf{x}_0 \in \mathscr{V}$, et vice versa.

De même, un vecteur $\mathbf{x}$ appartient par définition à la classe $\mathbf{x}_0 + \mathscr{P}$ si et seulement s'il s'écrit $\mathbf{x}_0 + \mathbf{a}$, $\mathbf{a} \in \mathscr{P}$. Algébriquement, cela signifie que, *primo*, la somme d'une solution fixe du système (9) et d'une solution arbitraire de (10) est solution de (9) et, *secundo*, on obtient de sorte toute solution de (9).

Diverses positions relatives des plans de l'espace $\mathscr{A}$ sont caractérisées algébriquement, disons, par les conditions à imposer aux rangs de certaines matrices et leurs sous-matrices. La question est suffisamment fastidieuse et difficile pour que nous passions outre.

* * *

L'existence de plans parallèles explique (ne fût-ce qu'en partie) le caractère encombrant de la théorie des plans de l'espace affine. Il est naturel que s'agissant de l'*espace projectif*, la théorie se simplifie légèrement (mais elle n'en reste pas moins ardue).

La définition générale d'un *espace projectif de dimension* n sur $\mathbb{K}$ quelconque a été donnée dans I.26. Selon cette définition, l'un des modèles de cet espace est l'ensemble de tous les sous-espaces de dimension 1 de l'espace vectoriel $\mathbb{K}^{n+1}$ de dimension $n+1$. On peut certes prendre au lieu de $\mathbb{K}^{n+1}$ tout espace vectoriel $\mathscr{V}^{n+1}$ de dimension $n+1$, donc tout espace affine $\mathscr{A} = \mathscr{A}^{n+1}$ de dimension $n+1$ muni de l'origine O. Dans le dernier cas, les points du modèle de l'espace projectif sont figurés par les droites de $\mathscr{A}$ passant par O, i.e. on retrouve le modèle « dans la gerbe » construit dans I.25 pour $n = 2$.

Pour fixer les idées, on considère le modèle $\mathbb{P}^{n+1}(\mathscr{V})$ dont les points sont des sous-espaces unidimensionnels de l'espace vectoriel $\mathscr{V}$ de dimension $n+1$.

Définition 5. On appelle *plan de dimension* r de l'espace projectif $\mathbb{P}^n(\mathscr{V})$ l'ensemble de tous les points de $\mathbb{R}^n(\mathscr{V})$ qui sont des sous-espaces unidimensionnels d'un espace $\mathscr{R} \subset \mathscr{V}$ de dimension $r+1$.

Ainsi, chaque plan de dimension r est par définition un espace projectif $\mathbb{P}^r(\mathscr{R})$ de dimension r.

On dit par abus de langage (mais en gagnant en brièveté) que les plans de $\mathbb{P}^n(\mathscr{V})$ sont des sous-espaces (de dimension supérieure d'une unité) de l'espace $\mathscr{V}$. Ainsi, nous avons par exemple identifié dans I.25 les droites du modèle $\mathscr{F}_0$ et les plans de l'espace affine $\mathscr{A}$ qui passent par le point O (i.e. les sous-espaces de dimension 2 de l'espace vectoriel attaché).

Dans I.25, nous avons expliqué en détail pour $n = 2$ que les *coordonnées projectives* $x^0 : x^1 : \ldots : x^n$ des points de $\mathbb{P}^n(\mathscr{V})$ sont définies par la donnée d'une base quelconque $\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_n$ de $\mathscr{V}$. Ce sont, pour chaque point M de $\mathbb{P}^n(\mathscr{V})$, les coordonnées dans cette base d'un vecteur $\mathbf{x} \in \mathscr{V}$ quelconque qui engendre M en tant que sous-espace de dimension 1 de $\mathscr{V}$. Autrement dit, $x^0 : x^1 : \ldots : x^n$ sont les coordonnées plückériennes du sous-espace M de dimension 1.

De façon plus générale, on définit (pour un système donné de coordonnées projectives) les *coordonnées projectives d'un plan* $\mathbb{P}^r(\mathscr{R}) \subset \mathbb{P}^n(\mathscr{V})$ *arbitraire* comme coordonnées plückériennes du sous-espace $\mathscr{R}$. Ainsi, ces coordonnées s'écrivent $p^{i_0 i_1 \ldots i_r}$ avec $i_0, i_1, \ldots, i_r = 0, 1, \ldots, n$, et sont assujetties aux conditions suivantes :

1) toute permutation $\sigma \in S_{r+1}$ vérifie l'égalité

$$p^{i_{\sigma(0)} \ldots i_{\sigma(r)}} = \varepsilon_\sigma p^{i_0 \ldots i_r}$$

(on suppose que σ opère sur les nombres $0, 1, \ldots, r$) ; il en découle que les coordonnées $p^{i_0 \ldots i_r}$ strictes sont au nombre de $\binom{n+1}{r+1}$; ce sont par exemple les coordonnées

$$p^{i_0 i_1 \ldots i_r} \quad \text{pour} \quad i_0 < i_1 < \ldots < i_r ;$$

2) on a pour n'importe quels indices $i_0, i_1, \ldots, i_r, j_0, j_1, \ldots, j_r$ la relation de Plücker

$$p^{[i_0 i_1 \ldots i_r} p^{j_0] j_1 \ldots j_r} = 0. \tag{11}$$

On montre que les relations (11) indépendantes sont exactement

$$N_{n,r} = \binom{n+1}{r+1} - (r+1)(n-r) - 1. \tag{12}$$

La démonstration directe exige une analyse combinatoire assez subtile. Au cours du semestre suivant, nous développerons la technique de calcul générale de ces constantes qui fournit le nombre (12) de façon triviale.

On peut se placer dans une géométrie dont les éléments de base (les « points») sont les plans r-dimensionnels de l'espace projectif de dimension n (ou, ce qui revient au même, les sous-espaces $r+1$-dimensionnels de l'espace vectoriel de dimension $n+1$). Analytiquement, on y arrive par recours aux coordonnées $p^{i_0 i_1 \ldots i_r}$, i.e. par identification des plans r-dimensionnels de l'espace projectif de dimension n et des points de coordonnées $p^{i_0 i_1 \ldots i_r}$, $i_0 < i_1 < \ldots$ $\ldots < i_r$, de l'espace projectif de dimension N, où $N = \binom{n+1}{r+1} - 1$.

Soit, par exemple, $r = n - 1$ (hyperplan). On a déjà dit que les coordonnées strictes sont dans ce cas $n+1$ coordonnées

$$p^{0 \ldots \hat{i} \ldots n} = (-1)^i q_i, \quad i = 0, 1, \ldots, n.$$

D'autre part, les hyperplans sont manifestement définis par une seule équation linéaire pour $x^0 : x^1 : \ldots : x^n$. Cela étant, on vérifie sans peine (vérifier !) que les nombres $q_0, q_1, \ldots, q_n$ sont justement les coefficients de cette équation. Dans I.24 et I.26, nous avons pris pour coordonnées des droites dans le plan ($n = 2$, $r = 1$) et des plans dans l'espace ($n = 3$, $r = 2$) les coefficients des équations des droites et des plans. Ainsi, les coordonnées plückériennes $p^{i_0 i_1 \ldots i_n}$ en sont une généralisation immédiate.

Le fait que $p^{i_0 \ldots i_n}$ ne sont assujettis pour $r = n - 1$ à aucune relation de Plücker non triviale (11) signifie que lorsqu'on représente les hyperplans par les points de l'espace projectif de dimension $\binom{n+1}{n} - 1$, on épuise les objets de l'espace. Comme $\binom{n+1}{n} - 1 = n$, il y a équivalence de la géométrie des hyperplans et de la géométrie des points. En tout cas, la première n'est pas plus ardue que la seconde.

Il en va autrement pour les droites dans l'espace de dimension 3 ($n = 3$, $r = 1$). Les coordonnées strictes $p^{i_0 i_1}$, $0 \leqslant i_0 < i_1 \leqslant 3$, sont au nombre de $\binom{4}{2} = 6$, si bien qu'elles définissent (par suite

de l'homogénéité) un point de l'espace KP^5 de dimension 5. Ces coordonnées vérifient de plus la relation

$$p^{01}p^{23} + p^{12}p^{03} + p^{02}p^{31} = 0$$

(voir plus haut ; les indices sont diminués d'une unité parce qu'ils prennent maintenant les valeurs de 0 à 3) qui détermine dans KP^5 une «hypersurface de degré 2». On voit donc que la géométrie des droites dans l'espace est équivalente à la géométrie des points d'une hypersurface « courbe» de l'espace de dimension 5. On ne s'étonne donc pas de ce que la première géométrie soit sensiblement plus compliquée que, disons, la géométrie des plans. C'est la raison pour laquelle nous l'avons en fait négligée pendant le premier semestre.

On conçoit que le cas de n et r quelconques est encore plus ardu. La variété de plans de dimension r de l'espace de dimension n est représentée par les points de l'espace de dimension $\binom{n+1}{r+1} - 1$ qui appartiennent à l'intersection de $N_{n,r}$ hypersurfaces de degré 2. Cette intersection porte le nom de *variété grassmannienne*. Bien que la variété grassmannienne fasse depuis des années l'objet de recherches intenses, sa géométrie renferme pas mal de points obscurs.

LEÇON 11

Fonctionnelles bilinéaires symétriques et antisymétriques. — Matrice d'une fonctionnelle bilinéaire symétrique. — Rang d'une fonctionnelle bilinéaire. — Fonctionnelles et formes quadratiques. — Théorème de Lagrange.

Par analogie avec les fonctionnelles antisymétriques, les *fonctionnelles multilinéaires symétriques* sont définies comme fonctionnelles B de $\mathsf{T}_p(\mathscr{V})$ (ou de $\mathsf{T}^p(\mathscr{V})$) telles que

$$\sigma B = B$$

quelle que soit la permutation $\sigma \in S_p$. La théorie correspondante s'avère très compliquée, si bien que le cas général de ces fonctionnelles n'est guère étudié. La seule exception est le cas $p = 2$ (fonctionnelles bilinéaires). Plaçons-nous donc dans ce cas et prenons, pour fixer les idées, les fonctionnelles de vecteurs (i.e. celles de $\mathsf{T}_2(\mathscr{V})$).

Le groupe S_2 étant à deux éléments (qui sont la permutation identique et la transposition (1 2)), *une fonctionnelle bilinéaire B est*

$$\left.\begin{array}{l}\textit{symétrique}\\ \textit{antisymétrique}\end{array}\right\}\ \textit{si et seulement si}\ \left\{\begin{array}{l}B(\mathbf{x},\mathbf{y}) = B(\mathbf{y},\mathbf{x}),\\ B(\mathbf{x},\mathbf{y}) = -B(\mathbf{y},\mathbf{x})\end{array}\right.$$

quels que soient les vecteurs $\mathbf{x}, \mathbf{y} \in \mathscr{V}$.

On sait que les fonctionnelles antisymétriques constituent le sous-espace $\Lambda_2(\mathscr{V}) \subset \mathsf{T}_2(\mathscr{V})$. De même, l'ensemble $\mathsf{S}_2(\mathscr{V})$ des fonctionnelles bilinéaires symétriques est un sous-espace de $\mathsf{T}_2(\mathscr{V})$.

Si le corps de base K est de caractéristique $= 2$, alors $\Lambda_2(\mathscr{V}) = \mathsf{S}_2(\mathscr{V})$. Dans le cas contraire, les sous-espaces $\Lambda_2(\mathscr{V})$ et $\mathsf{S}_2(\mathscr{V})$ sont évidemment disjoints:

$$\Lambda_2(\mathscr{V}) \cap \mathsf{S}_2(\mathscr{V}) = 0.$$

Dans la suite de l'exposé, cette hypothèse sera admise tacitement.

Proposition 1. *L'espace $\mathsf{T}_2(\mathscr{V})$ est somme directe de $\Lambda_2(\mathscr{V})$ et $\mathsf{S}_2(\mathscr{V})$, i.e. toute fonctionnelle bilinéaire B se décompose de façon unique en somme d'une fonctionnelle symétrique B_{sym} et d'une fonctionnelle antisymétrique B_{anti}:*

$$B = B_{\mathrm{sym}} + B_{\mathrm{anti}}. \tag{1}$$

Démonstration. L'unicité est garantie par l'hypothèse mentionnée. Pour prouver l'existence, il suffit de poser

$$B_{\text{sym}} = \frac{B+\sigma B}{2}, \quad B_{\text{anti}} = \frac{B-\sigma B}{2},$$

σ étant la transposition (1 2). □

On note que B_{anti} n'est rien autre que Alt B.

* * *

On sait (voir leçon 5) que la fonctionnelle B est bien définie dans une base donnée $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de l'espace $\mathscr{V}$ par sa matrice

$$\begin{pmatrix} b_{11} \ldots b_{1n} \\ \ldots\ldots \\ b_{n1} \ldots b_{nn} \end{pmatrix} \tag{2}$$

à éléments

$$b_{ij} = B(\mathbf{e}_i, \mathbf{e}_j), \quad i, j = 1, \ldots, n.$$

La valeur $B(\mathbf{x}, \mathbf{y})$ prise par B sur $\mathbf{x}, \mathbf{y} \in \mathscr{V}$ quelconques est une forme bilinéaire à coefficients b_{ij} de leurs coordonnées:

$$B(\mathbf{x}, \mathbf{y}) = b_{ij}x^i y^j.$$

Il en résulte de suite que la *fonctionnelle B est symétrique si et seulement s'il en est de même de sa matrice*, i.e. si

$$b_{ij} = b_{ji}$$

pour tout $i = 1, \ldots, n$ et tout $j = 1, \ldots, n$.

En effet, on a en particulier $B(\mathbf{e}_i, \mathbf{e}_j) = B(\mathbf{e}_j, \mathbf{e}_i)$, i.e. $b_{ij} = b_{ji}$, pour B symétrique. Inversement, si $b_{ij} = b_{ji}$, alors tout $\mathbf{x}$ et tout $\mathbf{y}$ vérifient l'égalité

$$B(\mathbf{y}, \mathbf{x}) = b_{ij}y^i x^j = b_{ji}y^j x^i = b_{ij}x^i y^j = B(\mathbf{x}, \mathbf{y}),$$

d'où la symétrie de B.

* * *

Comme pour toute fonctionnelle multilinéaire, on définit la notion du rang d'une fonctionnelle bilinéaire qui est égal au nombre minimal de covecteurs dont la fonctionnelle est combinaison tensorielle (voir leçon 7). S'agissant de B bilinéaire, on parle en outre du rang de sa matrice (2). Il serait bien tentant d'affirmer la coïncidence des deux notions. Il se trouve que cela n'a en général pas lieu.

Exemple. Soit $n = 2$ et $B = \mathbf{e}^1 \otimes \mathbf{e}^2$. La matrice de la fonctionnelle s'écrit

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

si bien qu'elle est de rang 1. Quant à la fonctionnelle B, elle s'exprime de manière tensorielle au moyen de deux covecteurs au moins (ne serait-ce que parce que toute fonctionnelle bilinéaire de rang 1 est symétrique).

Définition 1. Le rang de la matrice (2) constitue le *rang matriciel* de la fonctionnelle bilinéaire B.

On sait dès la leçon 5 que, dans un changement de base, la matrice de B est multipliée à gauche et à droite par les matrices C^{T} et C, avec C la matrice de passage. La multiplication par une matrice non dégénérée conserve le rang (leçon 2, prop. 2), ce qui montre le caractère intrinsèque de la définition 1.

Proposition 2. *Le rang matriciel r_{mat} d'une fonctionnelle bilinéaire B est au plus égal à son rang r:*

$$r_{\text{mat}} \leqslant r.$$

Démonstration. Soit $\xi^1, \ldots, \xi^r$ une base de l'espace associé à B, auquel cas

$$B = \sum_{i,\, j=1}^{r} b_{ij} \xi^i \otimes \xi^j,$$

b_{ij}, $i, j = 1, \ldots, r$, étant des nombres. Si l'on complète $\xi^1, \ldots, \xi^r$ de façon à former la base

$$\mathbf{e}^1 = \xi^1, \ldots, \mathbf{e}^r = \xi^r, \mathbf{e}^{r+1}, \ldots, \mathbf{e}^n$$

de l'espace $\mathscr{V}'$, dans la base correspondante $\mathbf{e}^i \otimes \mathbf{e}^j$, $i, j = 1, \ldots, n$, de l'espace $\mathsf{T}_2(\mathscr{V})$, on a

$$B = \sum_{i,\, j=1}^{r} b_{ij} \mathbf{e}^i \otimes \mathbf{e}^j$$

(voir leçon 5, prop. 3). Cela signifie que la matrice de la fonctionnelle par rapport à la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ s'écrit:

$$\begin{pmatrix} b_{11} & \ldots & b_{1r} & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{r1} & \ldots & b_{rr} & 0 & \ldots & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 \end{pmatrix}.$$

Aussi son rang est au plus égal à r. □

Le cas des fonctionnelles bilinéaires symétriques s'avère beaucoup plus simple.

Proposition 3. *Le rang matriciel d'une fonctionnelle bilinéaire symétrique B est identique au rang de B:*

$$r_{\text{mat}} = r.$$

Démonstration. Etant donnée la symétrie, tout covecteur associé est de la forme

$$\mathbf{x} \mapsto B(\mathbf{x}, \mathbf{a}), \tag{3}$$

$\mathbf{a}$ étant un vecteur quelconque. Tous ces covecteurs sont évidemment combinaisons linéaires des covecteurs

$$\mathbf{x} \mapsto B(\mathbf{x}, \mathbf{e}_i), \qquad i = 1, \ldots, n, \tag{4}$$

ce qui démontre que l'espace $\mathscr{R}$ associé à B est engendré par les covecteurs (4). Par conséquent, le rang $r = \dim \mathscr{R}$ est égal au rang de la famille de covecteurs (4). Il est clair que ces covecteurs ont pour coordonnées (dans la base $\mathbf{e}^1, \ldots, \mathbf{e}^n$) les colonnes de la matrice (2). Donc r est égal au rang de (2). □

La démonstration reste évidemment en vigueur pour les fonctionnelles bilinéaires antisymétriques, si bien que dans ce cas on a de même $r_{\text{mat}} = r$.

Dans les deux cas, les covecteurs (3) constituent manifestement un sous-espace. Par conséquent, les covecteurs associés ne font pas que engendrer l'espace associé à une fonctionnelle bilinéaire (anti) symétrique, ils en sont éléments.

* * *

Définition 2. Une fonctionnelle $Q : \mathbf{x} \mapsto Q(x) \in \mathbb{K}$ est dite *quadratique* s'il existe une fonctionnelle bilinéaire B telle que

$$Q(\mathbf{x}) = B(\mathbf{x}, \mathbf{x}) \tag{5}$$

pour tout vecteur $\mathbf{x} \in \mathscr{V}$.

On utilise la formule (1). Vu que $B_{\text{anti}}(\mathbf{x}, \mathbf{x}) = 0$, on estime sans restreindre la généralité que la fonctionnelle B de (5) est symétrique.

On vérifie de suite que B est alors définie de façon unique par Q, i.e. *la correspondance*

$$B \mapsto Q$$

est une bijection entre l'espace vectoriel $\mathsf{S}_2(\mathscr{V})$ des fonctionnelles bilinéaires symétriques et l'ensemble de toutes les fonctionnelles quadratiques sur $\mathscr{V}$ ($\mathbb{K}$ est toujours supposé de caractéristique $\neq 2$).

En effet, si $Q(\mathbf{x}) = B(\mathbf{x}, \mathbf{x})$ et si B est symétrique, alors

$$\frac{Q(\mathbf{x}+\mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y})}{2} = B(\mathbf{x}, \mathbf{y})$$

pour n'importe quels vecteurs $\mathbf{x}, \mathbf{y} \in \mathscr{V}$. □

Il est donc en principe indifférent de prendre une fonctionnelle bilinéaire symétrique ou une fonctionnelle quadratique car tout résultat relatif à la première peut être traduit en termes de la seconde, et vice versa. Nous préférons nous placer dans le cas de fonctionnelles

quadratiques en laissant la reformulation nécessaire au soin du lecteur.

Afin de simplifier les notations, la fonctionnelle bilinéaire symétrique associée à une fonctionnelle quadratique Q sera désignée par Q elle aussi. Son rang est par définition le *rang* de la fonctionnelle quadratique Q.

Dans chaque base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de l'espace $\mathscr{V}$, la fonctionnelle quadratique Q est définie par sa matrice

$$\begin{pmatrix} q_{11} & \cdots & q_{1n} \\ \cdot & \cdots & \cdot \\ q_{n1} & \cdots & q_{nn} \end{pmatrix} \tag{6}$$

à éléments

$$q_{ij} = Q(\mathbf{e}_i, \mathbf{e}_j), \qquad i, j = 1, \ldots, n.$$

La matrice (6) est carrée symétrique d'ordre n, et la correspondance

« la fonctionnelle » $\longmapsto$ « sa matrice »

est une bijection entre l'ensemble des fonctionnelles quadratiques sur $\mathscr{V}$ et l'ensemble des matrices symétriques d'ordre n à éléments dans K. Cette correspondance dépend de la base car la matrice (6) représentant Q dans une autre base est multipliée à gauche et à droite par C^{T} et C, avec C la matrice de passage. On fixe la base et on désigne la matrice (6) par Q.

Le rang de (6) est égal à celui de la fonctionnelle Q (proposition 3).

Tout vecteur $\mathbf{x} = x^i\mathbf{e}_i$ de $\mathscr{V}$ vérifie l'égalité

$$\begin{aligned} Q(\mathbf{x}) = q_{ij}x^ix^j = q_{11}(x^1)^2 + 2q_{12}x^1x^2 + \ldots + 2q_{1n}x^1x^n + \\ + q_{22}(x^2)^2 + \ldots\ 2q_{2n}x^2x^n + \\ \cdots\cdots\cdots\cdots\cdots \\ + q_{nn}(x^n)^2 \end{aligned}$$

ou, en notations matricielles,

$$Q(\mathbf{x}) = x^{\mathsf{T}}Qx,$$

où l'on a une fois de plus

$$x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix},$$

Q étant la matrice (6).

Définition 3. On dit qu'un polynôme $Q(x^1, \ldots, x^n)$ en les indéterminées $x^1, \ldots, x^n$ est une *forme quadratique* s'il est homogène (i.e. si tous ses éléments sont de même degré) de degré 2. Cf. I.14.

Toute forme quadratique s'écrit

$$\begin{aligned} Q(x^1, \ldots, x^n) = q_{ij}x^ix^j = \\ = q_{11}(x^1)^2 + 2q_{12}x^1x^2 + \ldots + 2q_{1n}x^1x^n + \\ + q_{22}(x^2)^2 + \ldots + 2q_{2n}x^2x^n + \\ \cdots\cdots\cdots\cdots\cdots \\ + q_{nn}(x^n)^2, \end{aligned}$$

si bien qu'elle est définie de façon unique par la matrice

$$\begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{pmatrix}$$

dite *associée* à la forme.

Ainsi, *la valeur* Q (x) *d'une fonctionnelle quadratique quelconque* Q *sur un vecteur* $\mathbf{x} \in \mathscr{V}$ *est donnée par la forme quadratique*

$$Q(\mathbf{x}) = Q(x^1, \ldots, x^n)$$

des coordonnées $x^1, \ldots, x^n$ *du vecteur* x.

Cela établit une bijection (dépendant de la base) entre les fonctionnelles quadratiques et les formes quadratiques.

Définition 4. Deux formes quadratiques sont dites *équivalentes* si elles représentent dans deux bases distinctes une même fonctionnelle quadratique.

On dit également que deux formes quadratiques sont équivalentes si les matrices associées Q_1, Q_2 vérifient l'égalité

$$Q_2 = C^{\mathrm{T}}Q_1C,$$

C étant une matrice non dégénérée.

Si l'on introduit les transformations linéaires homogènes

$$\begin{aligned} y^1 &= c_1^1x^1 + \ldots + c_n^1x^n, \\ &\cdots\cdots\cdots\cdots \\ y^n &= c_1^nx^1 + \ldots + c_n^nx^n \end{aligned} \tag{7}$$

à matrices non dégénérées

$$\begin{vmatrix} c_1^1 & \ldots & c_n^1 \\ \cdots & \cdots & \cdots \\ c_1^n & \ldots & c_n^n \end{vmatrix} \neq 0,$$

on dit des formes $Q_1(x^1, \ldots, x^n)$ et $Q_2(x^1, \ldots, x^n)$ qu'elles sont équivalentes à condition qu'il existe une transformation (7) telle qu'on obtienne Q_2 en notant $y^1, \ldots, y^n$ les variables de Q_1 et en utilisant les expressions (7).

* * *

Le produit scalaire introduit dans I.13 est un cas particulier de fonctionnelle bilinéaire symétrique caractérisé par l'axiome de positivité 15°. Si l'on fait abstraction de cet axiome, les espaces euclidiens font place aux espaces $\mathcal{V}$ dans lesquels on définit une fonctionnelle bilinéaire symétrique (ou, ce qui revient au même, une fonctionnelle quadratique) Q. On les appelle d'ordinaire *espaces pseudo-euclidiens* (avec des fois la restriction $\mathbb{K} = \mathbb{R}$).

Par analogie avec le cas euclidien, les vecteurs $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ seront dits *orthogonaux par rapport à* Q (ou *Q-orthogonaux*) si

$$Q(\mathbf{x}, \mathbf{y}) = 0.$$

Une question se pose : est-ce qu'il existe, pour les espaces pseudo-euclidiens, un analogue du procédé d'orthonormalisation de Gram-Schmidt (voir I.14) ? Du moment que la notion de famille orthonormée ne se généralise pas à ces espaces (comment orthonormer par exemple un vecteur $\mathbf{x} \neq \mathbf{0}$ pour lequel $Q(\mathbf{x}) = 0$?), on se demande naturellement s'il est possible de transformer une base quelconque en une base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *Q-orthogonale*, i.e. telle que

$$Q(\mathbf{e}_i, \mathbf{e}_j) = 0 \quad \text{pour} \quad i \neq j.$$

On a le

Théorème 1 (Lagrange). *Pour toute fonctionnelle quadratique Q sur $\mathcal{V}$ il existe une base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de l'espace $\mathcal{V}$ telle que*

$$Q(\mathbf{e}_i, \mathbf{e}_j) = 0 \quad \textit{pour} \quad i \neq j. \tag{8}$$

Démonstration. En plus de la démonstration du théorème, nous décrirons un algorithme pratique pour transformer une base quelconque de $\mathcal{V}$ en une base ayant la propriété (8).

Il s'agit de l'*algorithme de Lagrange* qui consiste à appliquer à tour de rôle trois transformations élémentaires dont l'une est *fondamentale* et les deux autres sont *auxiliaires*.

Transformation de Lagrange fondamentale. On l'applique à la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ si

$$q_{11} = Q(\mathbf{e}_1) \neq 0.$$

Par cette transformation $\mathbf{e}_1, \ldots, \mathbf{e}_n$ devient

$$\begin{aligned} \mathbf{e}'_1 &= \mathbf{e}_1, \\ \mathbf{e}'_2 &= -\frac{q_{12}}{q_{11}}\mathbf{e}_1 + \mathbf{e}_2, \\ &\ldots\ldots\ldots\ldots \\ \mathbf{e}'_n &= -\frac{q_{1n}}{q_{11}}\mathbf{e}_1 + \mathbf{e}_n \end{aligned} \tag{9}$$

telle que son premier vecteur est Q-orthogonal à tous les autres :

$$Q(\mathbf{e}'_1, \mathbf{e}'_i) = 0 \quad \text{pour} \quad i > 1.$$

En effet,

$$Q(\mathbf{e}_1', \mathbf{e}_i') = Q\left(\mathbf{e}_1, -\frac{q_{1i}}{q_{11}}\mathbf{e}_2 + \mathbf{e}_i\right) = -\frac{q_{1i}}{q_{11}}q_{11} + q_{1i} = 0.$$

Si $q_{22}' = Q(\mathbf{e}_2', \mathbf{e}_2') \neq 0$, on fait subir aux vecteurs $\mathbf{e}_2', \ldots, \mathbf{e}_n'$ (plus précisément à la restriction de Q au sous-espace $[\mathbf{e}_2', \ldots, \mathbf{e}_n']$) la même transformation et on obtient la base $\mathbf{e}_1'', \mathbf{e}_2'', \ldots, \mathbf{e}_n''$, avec $\mathbf{e}_1''$, $\mathbf{e}_2''$ Q-orthogonaux l'un à l'autre et aux vecteurs restants, et ainsi de suite.

Si le processus ne s'arrête pas jusqu'à la fin, i.e. si la transformation fondamentale est réalisée à chaque pas (tant qu'on n'épuise pas la base ou qu'on n'aboutit pas à une fonctionnelle nulle), le théorème 1 se trouve démontré. Ce cas est dit *régulier*.

Si la transformation fondamentale (9) devient impossible à une étape, on effectue des transformations auxiliaires fournissant une base susceptible de la transformation (9).

Première transformation auxiliaire. On l'utilise si $q_{11} = 0$ et s'il existe un indice i_0 tel que $q_{i_0 i_0} \neq 0$. Elle consiste à envoyer le i_0-ième vecteur de base à la place 1 :

$$\mathbf{e}_1' = \mathbf{e}_{i_0}, \quad \mathbf{e}_{i_0}' = \mathbf{e}_1,$$
$$\mathbf{e}_i' = \mathbf{e}_i, \quad \text{si } i \neq 1, i_0.$$

Dans la nouvelle base, on a évidemment $q_{11}' \neq 0$.

Seconde transformation auxiliaire. On y recourt si $q_{ii} = 0$ pour tous les $i = 1, \ldots, n$ et si la fonctionnelle Q n'est pas nulle si bien qu'on trouve i_0 et j_0 tels que $q_{i_0 j_0} \neq 0$. Si par exemple $q_{12} \neq 0$ (ce qui ne restreint manifestement pas la généralité de l'exposé), la transformation se définit comme suit :

$$\mathbf{e}_1' = \mathbf{e}_1 + \mathbf{e}_2,$$
$$\mathbf{e}_i' = \mathbf{e}_i \text{ si } i \geqslant 2.$$

Dans ce cas

$$q_{11}' = Q(\mathbf{e}_1') = Q(\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2) = 2q_{12} \neq 0,$$

ce qui permet la transformation fondamentale.

Le seul cas où aucune transformation mentionnée n'est possible est celui de tous les coefficients $q_{ij} = 0$, i.e. $Q = 0$. Il est évident que toute base étant visiblement Q-orthogonale dans ce cas ne nécessite pas de transformation.

Par conséquent, une succession déterminée des transformations décrites aboutit tôt ou tard à une base Q-orthogonale. □

Dans une base Q-orthogonale, la matrice de la forme Q est évidemment diagonale, i.e. elle s'écrit

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \tag{10}$$

donc

$$Q(\mathbf{x}) = \lambda_1 (x^1)^2 + \ldots + \lambda_n (x^n)^2$$

pour tout vecteur **x**. Aussi le théorème de Lagrange traduit en termes de formes quadratiques dit que *toute forme quadratique* $Q(x^1, \ldots, x^n)$ *est équivalente à la forme*

$$\lambda_1 (x^1)^2 + \ldots + \lambda_n (x^n)^2. \tag{11}$$

On dit de (11) qu'elle est *canonique*. Ainsi, *la transformation linéaire non dégénérée* (7) *réduit toute forme quadratique* $Q(x^1, \ldots, x^n)$ *à la forme canonique* (11).

La dernière affirmation qui porte également le nom de *théorème de Lagrange*, relève de l'algèbre seule, et elle ne laisse rien percer de son origine géométrique. On l'applique donc aux formes quadratiques intervenant dans tout problème (disons, en mécanique) qui est à priori étranger à la géométrie des fonctionnelles quadratiques.

Dans la pratique, on rend $Q(x^1, \ldots, x^n)$ canonique à l'aide de l'identité

$$Q(x^1, \ldots, x^n) = \frac{1}{q_{11}} (q_{11}x^1 + \ldots + q_{1n}x^n)^2 + Q'(x^2, \ldots, x^n),$$

où Q' ne renferme plus la variable x^1 (ce qu'il est immédiat de vérifier). L'identité ci-dessus correspond à la transformation de Lagrange fondamentale. Dans le cas non régulier, on procède en outre à un nouveau numérotage des variables et on utilise les transformations

$$y_1 = x_1 - x_2,$$
$$y_i = x_i \text{ si } i \geqslant 2.$$

Les coefficients de la forme (11) sont en partie (ou totalement) nuls. Il est clair que le nombre r de coefficients non nuls est égal au rang de la matrice (6) et, partant, à celui de la fonctionnelle Q. On permute, si besoin est, les éléments de base et on aboutit toujours à une forme dont les premiers coefficients $\lambda_1, \ldots, \lambda_r$ sont distincts de 0. Comme on omet les termes à coefficients nuls, *la forme canonique d'une forme quadratique de rang r* s'écrit finalement

$$\lambda_1 (x^1)^2 + \ldots + \lambda_r (x^r)^2, \text{ où } \lambda_1 \neq 0, \ldots, \lambda_r \neq 0.$$

LEÇON 12

Théorème de Jacobi. — Formes quadratiques sur le corps des nombres complexes et le corps des nombres réels. — Loi d'inertie. — Fonctionnelles et formes quadratiques définies positives.

On rappelle la définition d'une matrice triangulaire. Une matrice carrée est dite *triangulaire* (plus précisément triangulaire supérieure) si tous ses éléments au-dessous de la diagonale principale sont nuls. Son déterminant est évidemment le produit des éléments diagonaux. Aussi une matrice triangulaire est non dégénérée si et seulement si tous ses éléments diagonaux sont non nuls. On note l'intérêt particulier des matrices triangulaires ayant 1 pour éléments diagonaux. Nous les appellerons *matrices unitriangulaires*.

On trouve par un calcul direct que le produit de deux matrices (uni)triangulaires est une matrice (uni)triangulaire, ainsi que l'inverse d'une matrice (uni)triangulaire.

La matrice de l'étape fondamentale de l'algorithme de Lagrange étant unitriangulaire

$$\begin{pmatrix} 1 & -\frac{q_{12}}{q_{11}} & \dots & -\frac{q_{1n}}{q_{11}} \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 1 \end{pmatrix},$$

le passage à la base Q-orthogonale est donc associé dans le cas régulier à une matrice unitriangulaire.

Soient A une matrice carrée quelconque d'ordre n et $1 \leqslant k \leqslant n$. On biffe $n - k$ dernières lignes et $n - k$ dernières colonnes de A, et on obtient une matrice carrée d'orde k.

Définition 1. Cette matrice s'appelle *matrice principale d'ordre k extraite* de la matrice A, et son déterminant est le *mineur principal d'ordre k* de A.

Soient Q et Q' les matrices de la fonctionnelle quadratique Q dans les bases respectives $\mathbf{e}_1, \dots, \mathbf{e}_n$ et $\mathbf{e}'_1, \dots, \mathbf{e}'_n$, la matrice de passage C étant unitriangulaire. On a les affirmations évidentes:

a) la matrice principale C_k d'ordre k extraite de C est la matrice de passage de $\mathbf{e}_1, \dots, \mathbf{e}_k$ à $\mathbf{e}'_1, \dots, \mathbf{e}'_k$ du sous-espace $\mathcal{P}_k = [\mathbf{e}_1, \dots, \mathbf{e}_k] = [\mathbf{e}'_1, \dots, \mathbf{e}'_k]$;

b) la restriction $Q|_{\mathcal{P}_k}$ de la fonctionnelle Q au sous-espace $\mathcal{P}_k$ est une fonctionnelle quadratique dont la matrice dans la base $\mathbf{e}_1, \ldots, \mathbf{e}_k$ est la matrice principale Q_k d'ordre k extraite de Q et celle dans la base $\mathbf{e}'_1, \ldots, \mathbf{e}'_k$ la matrice principale Q'_k d'ordre k extraite de Q'.

Il en résulte

$$Q'_k = C_k^{\mathrm{T}} Q_k C_k$$

pour tout $k = 1, \ldots, n$. En passant aux déterminants, on a donc, par suite de $\det C_k^{\mathrm{T}} = \det C_k = 1$,

$$\det Q'_k = \det Q_k \tag{1}$$

quel que soit $k = 1, \ldots, n$.

En particulier,

$$\det Q_k = \lambda_1 \ldots \lambda_k$$

pour tout $k = 1, \ldots, n$ si

$$Q' = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Ainsi, on dit en termes de formes quadratiques que *si l'on est, pour une forme quadratique $Q(x^1, \ldots, x^n)$, dans le cas régulier, alors les coefficients $\lambda_1, \ldots, \lambda_n$ de sa forme canonique vérifient les relations*

$$\lambda_1 \ldots \lambda_k = D_k, \quad k = 1, \ldots, n, \tag{2}$$

avec D_k les mineurs principaux de la matrice de $Q(x^1, \ldots, x^n)$. □

On note que la transformation de Lagrange fondamentale donne chaque fois $\lambda \neq 0$ (une première transformation aboutit par exemple à $\lambda_1 = q_{11} \neq 0$). Dans le cas régulier, le processus s'arrête à un pas (disons, au r-ième pas) où on obtient $\lambda_r \equiv 0$ (si bien que $\lambda_{r+1}, \ldots, \lambda_n = 0$). Il en résulte, compte tenu de (2), que, *primo*, dans le cas régulier

$$D_1 \neq 0, \ldots, D_r \neq 0, \tag{3}$$

r étant le rang de la forme (de la fonctionnelle) et, *secundo*,

$$\lambda_1 = D_1, \quad \lambda_2 = \frac{D_2}{D_1}, \ldots, \lambda_r = \frac{D_r}{D_{r-1}}.$$

Inversement, si l'on suppose que la matrice de la forme quadratique satisfait à (3), r étant son rang, alors on applique la transformation de Lagrange fondamentale (car $q_{11} = D_1 \neq 0$). D'après les formules (1), les mineurs principaux de Q' ainsi obtenue coïncident avec ceux de la matrice Q, d'où la validité de (3) pour Q'. Mais D_2

de Q' est évidemment égal au produit $q'_{11}q'_{22}$ (où $q'_{11} = q_{11} = \lambda_1$), donc $q'_{22} \neq 0$. Par conséquent, la restriction de Q à $[e'_2, \ldots, e'_n]$ est susceptible de la transformation de Lagrange fondamentale, et ainsi de suite.

On a au bout de r pas une matrice de la forme

$$\begin{pmatrix} \lambda_1 & & & & 0 \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & \lambda_r & \\ 0 & & & & G \end{pmatrix}, \tag{4}$$

où $\lambda_1 \neq 0, \ldots, \lambda_r \neq 0$ et G est une matrice. Comme les matrices associées à une fonctionnelle dans des bases distinctes ont même rang r, la matrice (4) est de rang r, ce qui n'a évidemment lieu que si G a tous ses éléments nuls.

Ainsi, la matrice (4) est celle d'une forme quadratique réduite. Elle a été construite par les seules transformations fondamentales de l'algorithme de Lagrange, si bien que la forme initiale correspond au cas régulier.

On a donc démontré le

Théorème 1 (Jacobi). *Etant donnée une forme quadratique de rang r, on est dans le cas régulier si et seulement si r premiers mineurs principaux de sa matrice sont non nuls:*

$$D_1 \neq 0, \ldots, D_r \neq 0.$$

L'algorithme de Lagrange réduit cette forme à

$$D_1 (x^1)^2 + \frac{D_2}{D_1}(x^2)^2 + \ldots + \frac{D_r}{D_{r-1}}(x^r)^2. \ \square$$

Ce théorème rend souvent de grands services.

* * *

On simplifie la forme canonique

$$\lambda_1 (x^1)^2 + \ldots + \lambda_r (x^r)^2 \tag{5}$$

en fonction des propriétés arithmétiques du corps de base $\mathbb{K}$. Le cas le plus simple est $\mathbb{K} = \mathbb{C}$. Le processus

$$\begin{aligned} y^1 &= \sqrt{\lambda_1}\, x^1, \\ &\cdots\cdots \\ y^r &= \sqrt{\lambda_r}\, x^r, \\ y^{r+1} &= x^{r+1}, \\ &\cdots\cdots \\ y^n &= x^n \end{aligned}$$

aboutit alors à

$$(x^1)^2 + \ldots + (x^r)^2 \tag{6}$$

(on omet les accents des coordonnées).

Ainsi, on a la

Proposition 1. *Toute forme quadratique sur* $\mathbb{C}$ *(i.e. à coefficients dans* $\mathbb{C}$*) est ramenée par une transformation linéaire non dégénérée de variables (à coefficients dans* $\mathbb{C}$*) à* (6), *où r est le rang de la forme.* □

Autrement dit, toute forme quadratique $Q(x^1, \ldots, x^n)$ de rang r sur $\mathbb{C}$ s'écrit

$$\varphi_1(x)^2 + \ldots + \varphi_r(x)^2,$$

$\varphi_1(x), \ldots, \varphi_r(x)$ étant des formes linéaires linéairement indépendantes en $x^1, \ldots, x^n$.

Conséquence (théorème de classification des formes quadratiques sur $\mathbb{C}$**).** *Deux formes quadratiques sur* $\mathbb{C}$ *sont équivalentes si et seulement si elles sont de même rang.* □

Si $\mathbb{K} = \mathbb{R}$, on effectue les transformations

$$\begin{aligned} y^1 &= \sqrt{|\lambda_1|}\, x^1, \\ &\ldots\ldots\ldots \\ y^r &= \sqrt{|\lambda_r|}\, x^r, \\ y^{r+1} &= x^{r+1}, \\ &\ldots\ldots \\ y^n &= x^n \end{aligned}$$

qui ramènent (5) (après une permutation complémentaire éventuelle des coordonnées) à

$$(x^1)^2 + \ldots + (x^p)^2 - (x^{p+1})^2 - \ldots - (x^r)^2 \tag{7}$$

(on omet les accents des coordonnées), où r est le rang de la forme et p un nombre (tel que $0 \leqslant p \leqslant r$).

Cela démontre la

Proposition 2. *Toute forme quadratique sur* $\mathbb{R}$ *se ramène par une transformation linéaire non dégénérée de variables à la forme* (7), *r étant le rang de la forme et* $0 \leqslant p \leqslant r$. □

* * *

La proposition 2 entraîne de suite une question: est-ce qu'on réduit une forme quadratique donnée à deux formes (7) avec p différents? La réponse est négative, à savoir on a la

Proposition 3 (loi d'inertie des formes quadratiques). *Si deux formes*

$$(x^1)^2 + \ldots + (x^p)^2 - (x^{p+1})^2 - \ldots - (x^r)^2 \tag{8}$$

et

$$(y^1)^2 + \ldots + (y^q)^2 - (y^{q+1})^2 - \ldots - (y^r)^2 \tag{9}$$

sont équivalentes (sur $\mathbb{R}$*), alors* $p = q$.

Démonstration. L'équivalence de (8) et (9) signifie qu'elles représentent dans deux bases distinctes $\mathbf{e}_1, \ldots, \mathbf{e}_n$ et $\mathbf{f}_1, \ldots, \mathbf{f}_n$ une même fonctionnelle quadratique Q définie dans l'espace vectoriel $\mathcal{V}$ de dimension n. Soient $\mathcal{P}$ le sous-espace de $\mathcal{V}$ engendré par les vecteurs $\mathbf{e}_1, \ldots, \mathbf{e}_p$ et $\mathcal{Q}$ le sous-espace de $\mathcal{V}$ engendré par $\mathbf{f}_{q+1}, \ldots, \mathbf{f}_n$:

$$\mathcal{P} = [\mathbf{e}_1, \ldots, \mathbf{e}_p], \quad \mathcal{Q} = [\mathbf{f}_{q+1}, \ldots, \mathbf{f}_n].$$

Comme Q est représentée dans $\mathbf{e}_1, \ldots, \mathbf{e}_n$ par la forme (8), tout vecteur $\mathbf{x} \in \mathcal{P}$ non nul vérifie la relation

$$Q(\mathbf{x}) = (x^1)^2 + \ldots + (x^p)^2 > 0.$$

De même, on a pour tout $\mathbf{y} \in \mathcal{Q}$

$$Q(\mathbf{y}) = -(y^{q+1})^2 - \ldots - (y^r)^2 \leqslant 0.$$

Aussi $\mathcal{P} \cap \mathcal{Q} = 0$ et on a donc

$$\dim \mathcal{P} + \dim \mathcal{Q} \leqslant n$$

(conséquence 2 du théorème 1, leçon 1), i.e.

$$p + (n - q) \leqslant n,$$

ce qui équivaut à

$$p \leqslant q.$$

L'inégalité $q \leqslant p$ se démontre de façon analogue. Donc $p = q$. □

La proposition 3 justifie la

Définition 2. Le nombre p de « carrés positifs » de la forme réduite (7) s'appelle l'*indice d'inertie positif* de la forme quadratique (de la fonctionnelle quadratique) donnée, et le nombre $r - p$ de « carrés négatifs » en est *l'indice d'inertie négatif*.

La proposition 3 entraîne de suite la

Conséquence (théorème de classification des formes quadratiques sur $\mathbb{R}$). *Deux formes quadratiques sur $\mathbb{R}$ sont équivalentes si et seulement si elles ont même rang et mêmes indices d'inertie.*

* * *

S'agissant d'espaces vectoriels sur $\mathbb{R}$, on note l'intérêt tout particulier des fonctionnelles quadratiques Q telles que $Q(\mathbf{x}) > 0$ pour $\mathbf{x} \neq 0$. En effet, les fonctionnelles bilinéaires symétriques associées sont exactement tous les produits scalaires sur $\mathcal{V}$ (voir I.13, déf. 2).

Définition 3. Une fonctionnelle quadratique Q sur un espace vectoriel réel $\mathcal{V}$ est dite *définie positive* si $Q(\mathbf{x}) > 0$ pour tout vecteur $\mathbf{x} \neq 0$.

Une forme quadratique $Q(x^1, \ldots, x^n)$ est dite *définie positive* si elle représente dans une base une fonctionnelle définie positive, i.e. si $Q(x^1, \ldots, x^n) > 0$ pour $(x^1, \ldots, x^n) \neq (0, \ldots, 0)$.

Une matrice Q est dite *définie positive* si elle est la matrice d'une fonctionnelle (d'une forme) quadratique définie positive, i.e. si elle est la matrice des coefficients métriques rapportée à une base de l'espace euclidien (voir I.14).

Proposition 4. *Une fonctionnelle (une forme) quadratique est définie positive si et seulement si son rang et son indice d'inertie positif sont égaux à n:*

$$r = n, \qquad p = n.$$

Démonstration. Si $p = r = n$, la fonctionnelle Q s'exprime dans une base par la forme

$$(x^1)^2 + \ldots + (x^n)^2,$$

donc $Q(\mathbf{x}) = 0$ si et seulement si $x^1 = 0, \ldots, x^n = 0$, i.e. si et seulement si $\mathbf{x} = 0$.

Inversement, si ou bien $p < n$, ou bien $r < n$, la fonctionnelle Q est représentée dans une base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ par la forme

$$Q'(x^1, \ldots, x^{n-1}) + \varepsilon (x^n)^2,$$

où $Q'(x^1, \ldots, x^{n-1})$ est la forme quadratique des coordonnées $x^1, \ldots, x^{n-1}$ et $\varepsilon \leqslant 0$. Dans ce cas, $Q(\mathbf{e}_n) = \varepsilon \leqslant 0$ et, partant, Q n'est pas définie positive. □

La proposition 4 joue à condition d'effectuer au préalable la réduction à la forme canonique, si bien que son intérêt pratique est en général nul. On note par contre l'importance certaine de l'affirmation ci-dessous qui donne des conditions nécessaires et suffisantes de définie positivité d'une forme quadratique en fonction de sa matrice.

Proposition 5 (critère de Sylvester). *Une matrice Q est définie positive si et seulement si tous ses mineurs principaux*

$$q_{11} > 0, \quad \begin{vmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{vmatrix} > 0, \quad \ldots, \quad \begin{vmatrix} q_{11} & \cdots & q_{1n} \\ \cdot & \cdots & \cdot \\ q_{n1} & \cdots & q_{nn} \end{vmatrix} > 0.$$

Démonstration. Si tous les mineurs principaux de Q sont positifs (donc non nuls), le théorème 1 entraîne le cas régulier pour la forme quadratique de matrice Q, et la forme se ramène à

$$D_1 (x^1)^2 + \frac{D_2}{D_1} (x^2)^2 + \ldots + \frac{D_n}{D_{n-1}} (x^n)^2,$$

où $D_1 > 0, D_2 > 0, \ldots, D_n > 0$. Ainsi, $p = r = n$, si bien que la forme (donc la matrice) est définie positive.

Inversement, si la forme de matrice Q est définie positive, elle se réduit à la somme de n carrés, i.e. à une forme de matrice unité E. Aussi Q s'écrit

$$Q = C^{\top}C$$

(cf. I.14), avec C une matrice non dégénérée. Par conséquent,

$$\det Q = (\det C)^2 > 0,$$

i.e. *le déterminant d'une matrice définie positive est positif.*

D'autre part, on pose $x^{k+1}, \ldots, x^n = 0$ pour la forme quadratique $Q(x^1, \ldots, x^n)$ à n variables, il vient la forme quadratique

$$Q_k(x^1, \ldots, x^k) = Q(x^1, \ldots, x^k, 0, \ldots, 0)$$

à k variables $x^1, \ldots, x^k$ pour laquelle on dit évidemment que

a) si $Q(x^1, \ldots, x^n)$ est définie positive, il en est de même de $Q_k(x^1, \ldots, x^k)$;

b) la matrice de $Q_k(x^1, \ldots, x^k)$ est la matrice principale D_k d'ordre k extraite de la matrice de $Q(x^1, \ldots, x^n)$.

Conformément à la remarque ci-dessus, une matrice définie positive a donc tous ses mineurs principaux D_k, $k = 1, \ldots, n$, positifs. □

La proposition 5 répond en particulier à la question posée dans I.14 : quelles sont les conditions nécessaires et suffisantes pour qu'une matrice carrée soit la matrice des coefficients rapportée à une base de l'espace euclidien ?

LEÇON 13

Hypersurfaces du second degré dans l'espace projectif n-dimensionnel. — Hypersurfaces du second degré dans l'espace projectif complexe et réel-complexe. — Hypersurfaces du second degré dans l'espace affine n-dimensionnel. — Hypersurfaces du second degré dans l'espace affine complexe et réel-complexe.

Voyons ce que deviennent les résultats des leçons précédentes en ce qui concerne les formes quadratiques si l'on étudie les hypersurfaces du second degré dans l'espace projectif de dimension n.

Définition 1 (cf. I.25, déf. 2). On appelle *hypersurface du second degré* de l'espace projectif de dimension n (sur $\mathbb{K}$ quelconque de caractéristique $\neq 2$) l'ensemble des points dont les coordonnées projectives $x_0 : x_1 : \ldots : x_n$ vérifient l'équation

$$Q(x_0, x_1, \ldots, x_n) = 0,$$

$Q(x_0, x_1, \ldots, x_n)$ étant une forme quadratique des coordonnées $x_0, x_1, \ldots, x_n$. (Les indices inférieurs sont en l'occurrence plus commodes.)

Le théorème de Lagrange entraîne de suite le

Théorème 1 (réduction des équations des hypersurfaces du second degré de l'espace projectif n-dimensionnel sur $\mathbb{K}$ quelconque à la forme canonique). *Pour toute hypersurface du second degré de l'espace projectif de dimension n sur $\mathbb{K}$ de caractéristique $\neq 2$, il existe un système de coordonnées projectives $x_0 : x_1 : \ldots : x_n$ tel que l'équation de l'hypersurface dans ce système s'écrive*

$$\lambda_0 x_0^2 + \lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 = 0, \tag{1}$$

où $0 \leqslant r \leqslant n$ et $\lambda_1 \neq 0, \ldots, \lambda_r \neq 0$. □

Si $r = n$, on dit de (1) que c'est une *hypersurface ovale du second degré*.

On associe évidemment à tout plan Π_0 de dimension $k - 1$ de l'espace projectif et à tout point $M \notin \Pi_0$ de l'espace un seul plan $M\Pi_0$ de dimension k qui contient M et Π_0.

Définition 2. Une hypersurface de l'espace projectif de dimension n (sur $\mathbb{K}$ quelconque) s'appelle un *cylindre multiple d'ordre k* (ou un *cône multiple d'ordre k* car ces deux notions coïncident dans l'espace projectif) s'il existe un plan Π_0 de dimension $k - 1$ (appelé *plan méridien* du cylindre) tel qu'étant donné un point quelconque M, $M \notin \Pi_0$, de l'hypersurface, elle contienne le plan $M\Pi_0$ tout entier.

Chaque plan Π de dimension $n-k$ disjoint de Π_0 coupe le cylindre suivant une hypersurface de Π qui est par définition la *base du cylindre*. On dit également du cylindre concerné que c'est un cylindre *sur* sa base. Il est évident que chaque cylindre est réunion de tous les plans k-dimensionnels de la forme $M\Pi_0$, avec M un point quelconque de la base. En ce sens, la géométrie du cylindre se réduit à celle de sa base.

Fort de ces résultats, on considère l'hypersurface (1) pour $r < n$. Soit Π un plan de dimension r défini par $n-r$ équations $x_{r+1} = 0, \ldots, x_n = 0$. Dans ce plan, les nombres $x_0, x_1, \ldots, x_r$ sont les coordonnées projectives, et en ces coordonnées l'équation (1) détermine une hypersurface ovale du second degré. Soit, de plus, Π_0 un plan de dimension $n-r-1$ défini par $r+1$ équations $x_0 = 0$, $x_1 = 0, \ldots, x_r = 0$.

Le fait que toutes les fois que l'hypersurface (1) contient un point $(x_0^{(0)} : x_1^{(0)} : \ldots : x_n^{(0)})$, elle contient tous les points $(x_0^{(0)} : x_1^{(0)} : \ldots : x_r^{(0)} : x_{r+1} : \ldots : x_n)$, avec $x_{r+1}, \ldots, x_n$ des nombres arbitraires, signifie évidemment que (1) *est un cylindre multiple d'ordre* $n-r$ de plan méridien Π_0. La base du cylindre est l'hypersurface définie dans Π par l'équation (1).

Ainsi, on a le

Théorème 2 (classification des hypersurfaces du second degré de l'espace projectif de dimension n sur $\mathbb{K}$ quelconque). *Chaque hypersurface du second degré de l'espace projectif de dimension n sur $\mathbb{K}$ de caractéristique $\neq 2$ est soit une hypersurface ovale, soit un cylindre multiple d'ordre k, $1 \leqslant k \leqslant n$, sur une hypersurface ovale de l'espace projectif de dimension $n-k$.* □

L'espace projectif de dimension 1 est une droite, et toute « hypersurface » ovale de cet espace est un couple de points non confondus (ou un ensemble vide). Aussi le cylindre multiple d'ordre $n-1$ correspondant est un couple d'hyperplans distincts.

La situation se complique avec $k = n$. L'espace projectif de dimension 0 est un point, et toute hypersurface ovale de cet espace est un ensemble vide. D'autre part, l'équation $x_0^2 = 0$ définit dans l'espace projectif de dimension n $(n > 0)$ l'hyperplan « double » $x_0 = 0$. On dit donc, pour unifier la terminologie, que le cylindre multiple d'ordre n sur l'ensemble vide est un hyperplan de l'espace n-dimensionnel et que dans l'espace projectif de dimension 0 l'hypersurface ovale est l'ensemble vide « pris deux fois ».

Il y a intérêt à introduire la notion de *cylindre multiple d'ordre* 0 sur une hypersurface donnée, qu'on considère comme étant cette hypersurface. Dans ce cas, toute hypersurface du second degré de l'espace projectif de dimension n est un cylindre multiple d'ordre k, $0 \leqslant k \leqslant n$, sur une hypersurface ovale de l'espace de dimension $n-k$.

* * *

Si $K = \mathbb{C}$, tous les coefficients $\lambda_1, \ldots, \lambda_r$ de l'équation (1) sont supposés être égaux à 1, si bien qu'il existe pour tout r, $0 \leqslant r \leqslant n$, une hypersurface (1) unique. Le rang étant invariant, les hypersurfaces associées à r différents ne sont pas projectivement équivalentes. On vient de démontrer le

Théorème 3 (classification des hypersurfaces du second degré de l'espace projectif de dimension n sur $\mathbb{C}$). *Dans l'espace projectif complexe de dimension n il existe $n+1$ hypersurfaces du second degré qui ne sont pas projectivement équivalentes. Ce sont une hypersurface ovale et, pour tout r, $0 \leqslant r \leqslant n-1$, un cylindre multiple d'ordre $n-r$ sur une hypersurface ovale de l'espace r-dimensionnel.* □

On a établi dans I que s'agissant de $K = \mathbb{R}$, la situation géométrique n'est pas adéquate à la situation algébrique, si bien qu'on introduit les *espaces réels-complexes* (i.e. on se place dans le cas $(\mathbb{C}, \mathbb{R})$; cf. I.20).

Rien ne change du point de vue algébrique: toutes les transformations de coordonnées restent des transformations sur $\mathbb{R}$, et toutes les équations ont leurs coefficients réels.

Nous dirons d'une hypersurface du second degré du $(\mathbb{C}, \mathbb{R})$-espace projectif (ou affine; voir plus loin) qu'elle est *s-planaire* ($s \geqslant -1$) si elle ne contient pas de plan de dimension $s+1$, tandis qu'il passe par tout point réel de l'hypersurface au moins un plan (réel) de dimension s qu'elle contient tout entier. Dans l'espace de dimension 3 par exemple, l'hyperboloïde à deux nappes est 0-planaire et l'hyperboloïde à une nappe est 1-planaire. Une hypersurface est 1-planaire si et seulement si elle est dépourvue de points réels.

Dans le cas $(\mathbb{C}, \mathbb{R})$, l'équation (1) se ramène à

$$x_0^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_r^2 = 0, \quad 0 \leqslant r \leqslant n, \tag{2}$$

et on estime sans restreindre la généralité en multipliant par -1 que

$$-1 \leqslant p \leqslant \left[\frac{r+1}{2}\right] - 1,$$

avec $\left[\frac{r+1}{2}\right]$ la *partie entière* du nombre $\frac{r+1}{2}$ $\left(\text{si } \left[\frac{r+1}{2}\right] = m, \text{ alors } r = 2m \text{ ou } r = 2m-1\right)$.

Si $r = n$, l'hypersurface du second degré (2) est dite *non dégénérée*. On montre (le soin en est laissé au lecteur) que l'hypersurface (2) non dégénérée est p-planaire. Ainsi, elle n'a pas en particulier de point réel pour $p = -1$. On l'appelle *hypersurface ovale imaginaire du second degré*. Si $p = 0$, l'hypersurface (2) non dégénérée est dite *hypersurface ovale réelle du second degré*. Dans le cas $p > 1$ et $r = n$, on dit sans façon que (2) est une *hypersurface p-planaire non dégénérée du second degré*.

Comme la propriété d'être p-planaire est évidemment projectivement invariante, toutes les hypersurfaces (2) non dégénérées ne sont pas projectivement équivalentes.

Quand $r < n$, l'hypersurface (2) est un cylindre multiple d'ordre $n - r$ sur une hypersurface non dégénérée définie par (2) de l'espace de dimension r. Aussi la non-équivalence projective est propre à toutes les hypersurfaces (2).

Cela démontre le

Théorème 4 (classification des hypersurfaces du second degré de l'espace projectif réel-complexe de dimension n). *Dans l'espace projectif réel-complexe de dimension n ($n > 0$), il existe seulement $\left[\frac{n+1}{2}\right] + 1$ hypersurfaces non dégénérées projectivement non équivalentes du second degré qui ne sont pas des cylindres. Ce sont deux hypersurfaces ovales (réelle et imaginaire) et une hypersurface p-planaire par chaque $p = 1, \ldots, \left[\frac{n+1}{2}\right] - 1$ pour $n > 2$.*

Les hypersurfaces du second degré restantes sont des cylindres multiples d'ordre k, $1 \leqslant k \leqslant n$, sur des hypersurfaces non dégénérées de l'espace de dimension $n - k$ ou des hyperplans doubles (pour $k = n$). □

Des théorèmes analogues ont certes lieu dans l'*espace affine complété* déduit de l'espace projectif en choisissant un hyperplan en tant qu'hyperplan impropre. Dans cet espace, les hypersurfaces se distinguent de plus par leur position par rapport à l'hyperplan impropre. Les cylindres-cônes sont par exemple rangés en deux classes, la classe de cylindres dont le plan méridien Π_0 est tout entier dans l'hyperplan impropre et la classe de cylindres multiples d'ordre inférieur d'une unité sur les cônes dont Π_0 possède des points propres (si Π_0 est point propre, on a tout simplement des cônes). Aussi la classification des hypersurfaces du second degré de l'espace affine complété complexe a beau être triviale, elle est assez lourde.

L'espace affine complété privé d'hyperplan impropre devient l'espace affine. La classification des hypersurfaces du second degré de l'espace affine complexe dérive donc de celle des hypersurfaces du complété projectif, et les classes ainsi obtenues sont moins nombreuses. Notre souci de la clarté géométrique nous pousse néanmoins à l'établir de façon directe.

* * *

Soient $\mathcal{A}$ un espace affine de dimension n (sur $\mathcal{K}$ quelconque de caractéristique $\neq 2$) et $\mathcal{T}$ l'espace vectoriel associé.

Définition 3. On appelle *hypersurface du second degré* de $\mathcal{A}$ une partie de $\mathcal{A}$ formée de points dont les coordonnées affines $x_1, \ldots, x_n$ vérifient l'équation

$$F(x_1, \ldots, x_n) = 0,$$

$F(x_1, \ldots, x_n)$ étant un polynôme de degré 2 en $x_1, \ldots, x_n$. Cf. I.18, déf. 2.

On introduit le vecteur $\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$ (i.e. le rayon vecteur du point $M(x_1, \ldots, x_n)$) et on met l'équation $F(x_1, \ldots, x_n) = 0$ sous forme « vectorielle »

$$A(\mathbf{x}) + 2\alpha(\mathbf{x}) + a_{00} = 0, \tag{3}$$

avec A une fonctionnelle quadratique:

$$A(\mathbf{x}) = \sum_{i,j=1}^{n} a_{ij}x_ix_j,$$

α une fonctionnelle linéaire

$$\alpha(\mathbf{x}) = \sum_{i=1}^{n} a_{i0}x_i$$

et a_{00} un nombre. (Avec les conventions du premier semestre, on devrait écrire $n+1$ au lieu de l'indice 0, mais des considérations d'ordre technique nous obligent de changer de notations.)

On envoie l'origine O en un point O', ce qui donne pour chaque point $M \in \mathcal{A}$ le nouveau rayon vecteur $\mathbf{x}' = \overrightarrow{O'M}$ lié à $\mathbf{x} = \overrightarrow{OM}$ par la relation

$$\mathbf{x} = \mathbf{x}' + \mathbf{x}_0,$$

où $\mathbf{x}_0 = \overrightarrow{OO'}$. L'équation (3) est donc remplacée par

$$A(\mathbf{x}' + \mathbf{x}_0) + 2\alpha(\mathbf{x}' + \mathbf{x}_0') + a_{00} = 0,$$

i.e. par

$$A'(\mathbf{x}) + 2\alpha'(\mathbf{x}) + a_{00}' = 0,$$

on omet l'accent de $\mathbf{x}'$), où

$$\begin{aligned} A' &= A, \\ \alpha' &= \alpha_0 + \alpha, \\ a_{00}' &= A(\mathbf{x}_0) + 2\alpha(\mathbf{x}_0) + a_{00} \end{aligned} \tag{4}$$

(le symbole α_0 désigne le covecteur associé $\mathbf{x} \longrightarrow A(\mathbf{x}, \mathbf{x}_0)$).

Définition 4. Un point de rayon vecteur $\mathbf{x}_0 = x_1^{(0)}\mathbf{e}_1 + \ldots + x_n^{(0)}\mathbf{e}_n$ s'appelle *centre* de l'hypersurface (3) si

$$\alpha_0 + \alpha = 0,$$

i.e. si

$$\sum_{j=1}^{n} a_{ij}x_j^{(0)} + a_{i0} = 0, \quad i = 1, \ldots, n. \tag{5}$$

Les dernières relations forment un système de n équations à n inconnues $x_1^{(0)}, \ldots, x_n^{(0)}$. Si le système admet une solution unique, i.e. s'il existe un seul centre, l'hypersurface (3) est dite *à centre*.

Le système (5) a pour déterminant

$$\delta = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}, \tag{6}$$

déterminant de la matrice de la fonctionnelle A, si bien que le système admet pour $\delta \neq 0$ par la règle de Cramer une solution unique. Si $\delta = 0$, le rang r de la matrice de (5) est inférieur à n, ce qui entraîne (par Kronecker-Capelli), ou bien l'incompatibilité (absence de centres), ce qui a lieu quand la matrice

$$\begin{pmatrix} a_{10} & a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n0} & a_{n1} & \cdots & a_{nn} \end{pmatrix} \tag{7}$$

est de rang $r + 1$, ou bien la propriété, pour (5), de définir dans $\mathcal{A}$ un plan (*plan des centres*) de dimension $n - r$, ce qui a lieu pour le rang r de la matrice (7).

Ainsi, *l'hypersurface* (3) *est une hypersurface à centre si et seulement si* $\delta \neq 0$. □

Si l'hypersurface (3) a au moins un centre, on obtient par identification de l'origine et du centre

$$A(\mathbf{x}) + a_{00} = 0.$$

Cela étant, on ne restreint en rien la généralité si l'on divise l'équation par a_{00}, $a_{00} \neq 0$. On choisit de plus (par Lagrange) une base du repère telle qu'on ait

$$A(\mathbf{x}) = \lambda_1 x_1^2 + \ldots + \lambda_r x_r^2, \tag{8}$$

où $\lambda_1 \neq 0, \ldots, \lambda_r \neq 0$. Ainsi, *il existe pour une hypersurface du second degré ayant un centre un système des coordonnées affines* $x_1, \ldots, x_n$ *en lesquelles elle a pour équation*

$$\lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 = \varepsilon,$$

où $\lambda_1 \neq 0, \ldots, \lambda_r \neq 0$ *et* $\varepsilon = 0, 1$. □

Si $r = n$, l'hypersurface est à centre, et elle est pour $r < n$ un cylindre multiple d'ordre $n - r$ sur une hypersurface à centre de l'espace de dimension r.

On suppose maintenant que l'hypersurface (3) n'a pas de centre (on note que cela n'est possible que pour $n > 1$). Cela signifie que le covecteur α du dual $\mathcal{V}''$ n'est pas de la forme $-\alpha_0$, i.e. il n'est pas associé à la fonctionnelle bilinéaire A et n'appartient donc pas à l'espace associé $\mathcal{H}$.

On ramène la fonctionnelle quadratique A à la forme (8), i.e. on trouve dans l'espace $\mathscr{H}$ associé à la fonctionnelle bilinéaire A attachée une base $\mathbf{e}^1, \ldots, \mathbf{e}^r$ telle que

$$A = \lambda_1 \mathbf{e}^1 \otimes \mathbf{e}^1 + \ldots + \lambda_r \mathbf{e}^r \otimes \mathbf{e}^r.$$

(On obtient la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de $\mathscr{T}$ dans laquelle la fonctionnelle quadratique A est représentée par la formule (8), en complétant $\mathbf{e}^1, \ldots, \mathbf{e}^r$ en la base $\mathbf{e}^1, \ldots, \mathbf{e}^n$ de $\mathscr{T}'$ et en passant à la base duale.)

Etant donné $\alpha \notin \mathscr{H}$, on choisit évidemment $\mathbf{e}^1, \ldots, \mathbf{e}^n$ telle qu'on vérifie l'égalité $\mathbf{e}^{r+1} = -\alpha$, i.e. telle que $\alpha(\mathbf{x}) = -x_{r+1}$ pour tout vecteur $\mathbf{x} \in \mathscr{T}$.

Dans cette base, l'équation (3) s'écrit pour tout point origine O:

$$\lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 - 2x_{r+1} + a_{00} = 0. \tag{9}$$

Si l'on transfère l'origine O en le point de coordonnées

$$\Big(\underbrace{0, \ldots, 0}_{r \text{ fois}}, \frac{a_{00}}{2}, 0, \ldots, 0\Big),$$

on obtient manifestement (voir dernière formule (4)) une équation de la forme (9), où $a_{00} = 0$.

On a donc démontré le

Théorème 5 (réduction des équations des hypersurfaces du second degré de l'espace affine de dimension n sur $\mathbb{K}$ quelconque à la forme canonique). *Pour toute hypersurface du second degré de l'espace affine de dimension n sur $\mathbb{K}$ de caractéristique $\neq 2$, il existe un système de coordonnées affines en lesquelles son équation est, ou bien*

$$\lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 = \varepsilon, \tag{I}$$

où $1 \leqslant r \leqslant n$ et $\varepsilon = 0$ ou 1, ou bien

$$\lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 = 2x_{r+1} \tag{II}$$

(pour $n > 1$ seulement), où $1 \leqslant r \leqslant n - 1$, dans les deux cas $\lambda_1 \neq 0, \ldots, \lambda_r \neq 0$. □

Si $r = n$ et $\varepsilon = 1$, l'hypersurface (I) s'appelle *une hypersurface ovale du second degré*. Si $r = n$ et $\varepsilon = 0$, c'est *un cône du second degré* qui est un cône sur une hypersurface ovale de l'espace de dimension $n - 1$. Quand $r < n$, (I) constitue un cylindre multiple d'ordre $n - r$ dont la base est, ou bien une hypersurface ovale (cas $\varepsilon = 1$), ou bien un cône du second degré (cas $\varepsilon = 0$) de l'espace de dimension r.

L'hypersurface (II) s'appelle pour $r = n - 1$ un *paraboloïde*. Si $r < n - 1$, c'est un cylindre multiple d'ordre $n - r - 1$ sur un paraboloïde de l'espace de dimension $r + 1$.

Ainsi, on a le

Théorème 6 (classification des hypersurfaces du second degré de l'espace affine de dimension n sur $\mathbb{K}$ quelconque). *Chaque hypersurface du second degré de l'espace affine de dimension n sur $\mathbb{K}$ de caractéristique $\neq 2$ est*

a) *ou bien une hypersurface ovale,*

b) *ou bien un cône,*

c) *ou bien un paraboloïde (pour $n > 1$),*

d) *ou bien un cylindre multiple d'ordre k, $1 \leqslant k \leqslant n-1$, sur une hypersurface* a), b) *ou* c) *de l'espace affine de dimension $n-k$.*

Cela étant, deux hypersurfaces de types différents ne sont pas affinement équivalentes.

La dernière affirmation découle des résultats suivants:

1) les hypersurfaces b) possèdent un *sommet* (i.e. un point tel que la droite contenant ce point et un point quelconque de l'hypersurface soit tout entière sur l'hypersurface) et les hypersurfaces a) en sont dépourvues;

2) les hypersurfaces a) et b) possèdent un *centre de symétrie* et les hypersurfaces c) n'en ont pas;

3) les hypersurfaces d) sont des cylindres, ce qui n'est pas vrai de a), b) et c). □

* * *

Lorsque $\mathbb{K} = \mathbb{C}$, les classes a), b) et c) renferment chacune, à l'équivalence affine près, une seule hypersurface du second degré, si bien qu'on a le

Théorème 7 (classification des hypersurfaces du second degré de l'espace affine de dimension n sur $\mathbb{C}$). *Dans l'espace affine de dimension n sur $\mathbb{C}$, il existe pour $n = 1$ deux hypersurfaces du second degré qui ne sont pas affinement équivalentes. Ce sont une hypersurface ovale formée de deux points non confondus et un cône du second degré réduit à deux points confondus. Si $n > 1$, il en existe trois qui ne sont pas des cylindres: une hypersurface ovale, un cône du second degré et un paraboloïde, les hypersurfaces restantes étant des cylindres multiples d'ordre k, $1 \leqslant k \leqslant n-1$, sur trois (ou deux si $k = n-1$) hypersurfaces mentionnées de l'espace affine de dimension $n-k$.* □

Si $\mathbb{K} = \mathbb{R}$ (cas $(\mathbb{C}, \mathbb{R})$), l'équation (I) se ramène à

$$(\mathrm{I}') \qquad x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_r^2 = \varepsilon,$$

où $\varepsilon = -1$, 0 ou 1 et $1 \leqslant r \leqslant n$, et l'équation (II) à

$$(\mathrm{II}') \qquad x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_r^2 = 2x_{r+1},$$

avec $1 \leqslant r \leqslant n-1$, et le fait de multiplier par -1 (et de changer le signe de la coordonnées x_{r+1}) permet d'estimer sans restreindre la généralité que $0 \leqslant p \leqslant \frac{n}{2}$ dans les deux cas (et que de plus $\varepsilon \neq -1$ pour $p = \frac{n}{2}$, donc pour n pair dans le cas (I')).

Quand $r = n$, l'hypersurface (I') est par définition une *hypersurface non dégénérée du second degré*. Si $\varepsilon \neq 0$ et $p = 0$, une hypersurface non dégénérée est un *ellipsoïde* qui est *imaginaire* ou *réel* selon que $\varepsilon > 0$ ou $\varepsilon < 0$. Ce sont pour $n = 2$ une ellipse imaginaire et une ellipse réelle respectivement. Lorsque $n = 1$, on a les couples de points imaginaires ou réels.

Une hypersurface non dégénérée du second degré est pour $\varepsilon = 0$ un *cône du second degré*. Un cône du second degré contient pour $p = 0$ un point réel unique, ce qui légitime son nom de *cône imaginaire du second degré*.

Une hypersurface non dégénérée du second degré constitue pour $\varepsilon \neq 0$, $1 \leqslant p \leqslant \frac{n}{2}$ un *ε-hyperboloïde*.

Lorsque $n = 2$, il existe un seul hyperboloïde (une hyperbole) et deux cônes représentés par les couples de droites sécantes imaginaires et réelles. Le cas $n = 1$ ne correspond à aucun hyperboloïde, mais on compte un cône réduit à deux points confondus.

Comme dans l'espace projectif, une hypersurface du second degré de l'espace affine réel-complexe est dite *s-planaire* s'il passe par tout point réel de l'hypersurface au moins un plan de dimension s qu'elle contient tout entier, tandis que tout plan de dimension $s + 1$ lui est extérieur.

On montre (le lecteur est prié de le faire) que tout ε-hyperboloïde est s-planaire, avec $s = p - 1$ pour $\varepsilon = 1$ et $s = p$ pour $\varepsilon = -1$, et que chaque cône du second degré est p-planaire.

Dans le cas $r < n$, les hypersurfaces (I') sont des cylindres multiples d'ordre $n - r$ sur des hypersurfaces non dégénérées de l'espace de dimension r.

Si $r = n - 1$, l'hypersurface (II') s'appelle un *paraboloïde* qui est *elliptique* pour $p = 0$ (si $n = 2$, on a une parabole) et *hyperbolique* pour $1 \leqslant p \leqslant \frac{n}{2}$. On montre (le faire!) que tout paraboloïde est p-planaire.

Si $r < n - 1$, l'hypersurface (II') constitue un cylindre multiple d'ordre $n - r - 1$ sur un paraboloïde de l'espace r-dimensionnel.

On démontre comme pour $\mathbb{K} = \mathbb{C}$ que les ellipsoïdes, les hyperboloïdes, les cônes, les paraboloïdes et les cylindres ne sont pas affinement équivalents. S'agissant des paraboloïdes, cela tient à ce qu'ils sont p-planaires pour p différents. L'explication est la même en ce qui concerne les cônes et les ε-hyperboloïdes à ε identique. Il est évident que l'ellipsoïde réel et l'ellipsoïde imaginaire ne sont pas affinement équivalents entre eux ni aux ε-hyperboloïdes, à l'exception peut-être du 1-hyperboloïde pour $p = 1$ (i.e. du 1-hyperboloïde 0-planaire). Or les plans de dimension 2 coupent ce dernier entre autres suivant des hyperboles, ce qui n'a pas lieu pour l'ellip-

soïde. Aussi il y a non-équivalence affine de l'ellipsoïde et du 1-hyperboloïde 0-planaire. Les plans qui coupent le 1-hyperboloïde s-planaire (associé à la valeur $p = s + 1$) suivant l'ellipsoïde imaginaire (i.e. qui ne le rencontrent pas dans le domaine réel) sont pour $s > 1$ de dimension au plus égale à $n - s - 1 = n - p$ (démontrer !), tandis que s'agissant du -1-hyperboloïde s-planaire (pour lequel $p = s$), elle est $s = p$ (démontrer !). L'égalité $p = n - p$ étant impossible (car le cas $\varepsilon = -1$ est pour $n = 2p$ exclu par hypothèse), les ± 1-hyperboloïdes s-planaires ne sont pas affinement équivalents.

Cela démontre le

Théorème 8 (classification des hypersurfaces du second degré de l'espace affine réel-complexe de dimension n). *Dans l'espace affine réel-complexe de dimension n, les seules hypersurfaces du second degré ayant la propriété de ne pas être affinement équivalentes qui ne sont pas des cylindres sont*

a) *deux ellipsoïdes (réel et imaginaire)*;

b) *un 1-hyperboloïde s-planaire pour tout $s = 0, 1, \ldots, \left[\frac{n}{2}\right] - 1$*;

c) *un -1-hyperboloïde s-planaire pour tout $s = 1, \ldots, m$, avec $m = \frac{n}{2} - 1$ si n est pair et $m = \frac{n-1}{2}$ dans le cas contraire*;

d) *un cône du second degré p-planaire pour tout $p = 0, 1, \ldots, \left[\frac{n}{2}\right]$ (c'est un cône imaginaire pour $p = 0$)*;

e) *un paraboloïde p-planaire pour tout $p = 0, \ldots, \left[\frac{n}{2}\right]$ (si $p = 0$, on a un paraboloïde elliptique, et si $p = 1, \ldots, \left[\frac{n}{2}\right]$, ce sont des paraboloïdes hyperboliques).*

Les hypersurfaces du second degré restantes sont des cylindres multiples d'ordre k, $1 \leqslant k \leqslant n - 1$, sur les hypersurfaces énumérées de l'espace affine de dimension $n - k$. □

LEÇON 14

Algèbre des opérateurs linéaires. — Opérateurs et fonctionnelles bilinéaires mixtes. — Opérateurs linéaires et matrices. — Opérateurs inversibles. — Opérateur adjoint. — Alternative de Fredholm. — Sous-espaces invariants et opérateurs induits.

Revenons à la théorie des espaces vectoriels et plaçons-nous dans le cas de fonctionnelles bilinéaires mixtes B: $\mathbf{x}, \xi \mapsto B(\mathbf{x}, \xi)$, $\mathbf{x} \in \mathcal{T}$, $\xi \in \mathcal{T}'$ (dont l'étude est à peine entamée dans la leçon 5). Il se trouve que les fonctionnelles mixtes sont étroitement liées à des homomorphismes (voir leçon 3, déf. 5) tels que $\mathcal{W} = \mathcal{T}$.

Définition 1. Les homomorphismes de $\mathcal{T}$ dans $\mathcal{T}$ s'appellent les *opérateurs linéaires sur* $\mathcal{T}$.

Ainsi, l'application

$$\mathbf{A} : \mathcal{T} \to \mathcal{T} \tag{1}$$

est un opérateur linéaire si

$$\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y}$$

et

$$\mathbf{A}(k\mathbf{x}) = k\mathbf{A}\mathbf{x}$$

quels que soient les vecteurs $\mathbf{x}, \mathbf{y} \in \mathcal{T}$ et le nombre $k \in \mathbb{K}$.

La *somme* $\mathbf{A} + \mathbf{B}$ de deux opérateurs linéaires $\mathbf{A}$ et $\mathbf{B}$ et le *produit* $k\mathbf{A}$ d'un opérateur linéaire $\mathbf{A}$ par un nombre $k \in \mathbb{K}$ sont définis de façon usuelle par les formules

$$(\mathbf{A} + \mathbf{B})\,\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x},$$
$$(k\mathbf{A})\,\mathbf{x} = k\,(\mathbf{A}\mathbf{x})$$

et constituent évidemment des opérateurs linéaires. Il est immédiat de vérifier que l'*ensemble* Op ($\mathcal{T}$) *de tous les opérateurs linéaires sur* $\mathcal{T}$ *est un espace vectoriel* pour ces opérations.

L'élément nul de cet espace est l'*opérateur* $\mathbf{O}$ s'exerçant par la formule

$$\mathbf{O}(\mathbf{x}) = \mathbf{0}.$$

On définit de même l'*opération de multiplication* de deux opérateurs $\mathbf{A}, \mathbf{B} \mapsto \mathbf{AB}$. Comme c'est l'usage pour les applications, on prend pour *produit* $\mathbf{AB}$ la composée $\mathbf{A} \circ \mathbf{B}$ des opérateurs $\mathbf{A}$ et $\mathbf{B}$.

Ainsi,

$$(\mathbf{AB})\,\mathbf{x} = \mathbf{A}\,(\mathbf{Bx})$$

pour tout vecteur $\mathbf{x} \in \mathscr{V}$. L'opérateur **AB** est évidemment linéaire.

On montre par un calcul banal l'associativité de la multiplication d'opérateurs:

$$(\mathbf{AB})\,\mathbf{C} = \mathbf{A}\,(\mathbf{BC})$$

(si bien qu'on omet les parenthèses dans le produit d'opérateurs en nombre quelconque) et sa distributivité par rapport à l'addition:

$$\mathbf{A}\,(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC},$$

i.e. *l'ensemble* Op ($\mathscr{V}$) *est un anneau.*

Cet anneau a un élément unité qui est l'opérateur identique $\mathbf{E}: \mathscr{V} \to \mathscr{V}$ laissant invariant chaque vecteur $\mathbf{x} \in \mathscr{V}$:

$$\mathbf{Ex} = \mathbf{x}.$$

En général, $\mathbf{AB} \neq \mathbf{BA}$, d'où la non-commutativité de l'anneau Op ($\mathscr{V}$) (pour $n > 1$).

Il y a entre le produit de deux opérateurs et le produit d'un opérateur par un nombre $k \in \mathbb{K}$ la relation

$$(k\mathbf{A})\,\mathbf{B} = \mathbf{A}\,(k\mathbf{B}) = k\,(\mathbf{AB}) \tag{2}$$

démontrée par un calcul trivial.

Les anneaux qui sont des espaces vectoriels et remplissent les égalités (2) sont dits des *algèbres*. Fort des résultats ci-dessus, on peut dire de l'ensemble Op ($\mathscr{V}$) qu'*il constitue une algèbre*. □

La relation (2) entraîne en particulier

$$(k\mathbf{E})\,\mathbf{A} = \mathbf{A}\,(k\mathbf{E})$$

pour tout opérateur **A**. Ainsi, les opérateurs de la forme $k\mathbf{E}$ appelés *opérateurs scalaires* permutent avec tous les opérateurs.

Il se trouve que cette propriété caractérise les opérateurs scalaires, i.e. tout opérateur permutable avec chaque opérateur de Op ($\mathscr{V}$) est scalaire (le montrer !). On dit donc que l'algèbre Op ($\mathscr{V}$) est non commutative au plus haut degré (permis par la structure algébrique).

* * *

Chaque opérateur **A** définit par la formule

$$A\,(\mathbf{x}, \xi) = \xi\,(\mathbf{Ax})$$

une fonctionnelle bilinéaire mixte $A \in \mathrm{T}^1_1\,(\mathscr{V})$. Inversement, étant donnée une fonctionnelle bilinéaire mixte A quelconque, la correspondance entre un vecteur $\mathbf{x} \in \mathscr{V}$ arbitraire et le covecteur associé

$$\mathbf{Ax}: \xi \mapsto A\,(\mathbf{x}, \xi)$$

de l'espace $\mathscr{V}''$ (i.e. un vecteur de $\mathscr{V}'$ par identification $(\mathscr{V}')' = \mathscr{V}$) est un opérateur linéaire $\mathbf{A} \in \mathrm{Op}\,(\mathscr{V})$. Les applications $\mathbf{A} \mapsto A$ et $A \mapsto \mathbf{A}$ étant évidemment réciproques l'une de l'autre sont bijectives. Elles transforment manifestement toute somme en une somme et tout produit par un nombre en un produit par le même nombre, si bien que ce sont des isomorphismes. Ainsi, *les espaces vectoriels* $\mathrm{Op}\,(\mathscr{V})$ *et* $\mathrm{T}^1_1\,(\mathscr{V})$ *sont canoniquement isomorphes.* □

Nous identifierons de règle un opérateur et la fonctionnelle bilinéaire correspondante.

* * *

Soit, dans l'espace $\mathscr{V}$, une base $\mathbf{e}_1, \ldots, \mathbf{e}_n$, auquel cas on a pour tout vecteur $\mathbf{x} = x^1\mathbf{e}_1 + \ldots + x^n\mathbf{e}_n$

$$\mathbf{A}\mathbf{x} = x^1\mathbf{a}_1 + \ldots + x^n\mathbf{a}_n, \tag{3}$$

où $\mathbf{a}_1 = \mathbf{A}\mathbf{e}_1, \ldots, \mathbf{a}_n = \mathbf{A}\mathbf{e}_n$. Inversement, la formule (3) définit de façon unique pour toute famille de vecteurs $\mathbf{a}_1, \ldots, \mathbf{a}_n$ un opérateur linéaire $\mathbf{A}$ tel que $\mathbf{a}_1 = \mathbf{A}\mathbf{e}_1, \ldots, \mathbf{a}_n = \mathbf{A}\mathbf{e}_n$. On dit qu'avec une base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ fixe, *on a une correspondance bijective entre les opérateurs* $\mathbf{A} \in \mathrm{Op}\,(\mathscr{V})$ *et les familles de n vecteurs* $\mathbf{a}_1, \ldots, \mathbf{a}_n$. □

Il correspond à chaque famille une matrice carrée dont les *colonnes* ont pour éléments les coordonnées des vecteurs $\mathbf{a}_1, \ldots, \mathbf{a}_n$ (dans la même base $\mathbf{e}_1, \ldots, \mathbf{e}_n$):

$$A = \begin{pmatrix} a^1_1 & \ldots & a^1_n \\ \cdot & \cdot\;\cdot\;\cdot & \cdot \\ a^n_1 & \ldots & a^n_n \end{pmatrix}. \tag{4}$$

Puisqu'on établit de sorte une correspondance bijective entre les matrices et les familles $\mathbf{a}_1, \ldots, \mathbf{a}_n$, il en est de même des opérateurs et des matrices carrées d'ordre n. On vérifie automatiquement qu'il s'agit d'un isomorphisme (car dans cette transformation toute somme devient une somme et tout produit par un nombre change en un produit par le même nombre).

Cela démontre la

Proposition 1. *Le choix d'une base* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *de l'espace vectoriel* $\mathscr{V}$ *de dimension n sur* $\mathbb{K}$ *établit un isomorphisme entre l'algèbre des opérateurs* $\mathrm{Op}\,(\mathscr{V})$ *et l'algèbre des matrices carrées d'ordre n sur* $\mathbb{K}$. □

Cela étant, à un opérateur $\mathbf{A}$ correspond par cet isomorphisme une matrice A à colonnes formées de coordonnées des vecteurs $\mathbf{A}\mathbf{e}_1, \ldots, \mathbf{A}\mathbf{e}_n$ dans la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

Définition 2. La matrice A s'appelle la *matrice de l'opérateur* $\mathbf{A}$ *dans la base* $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

Comme $a^j_i = \mathbf{e}^j\,(\mathbf{A}\mathbf{e}_i)$, la matrice A est également la matrice de la fonctionnelle bilinéaire mixte A. Il en résulte (voir leçon 5) que la matrice $A' = (a^{j'}_{i'})$ de l'opérateur $\mathbf{A}$ dans toute autre base

$\mathbf{e}_{1'}, \ldots, \mathbf{e}_{n'}$ se définit par

$$A' = C^{-1}AC, \tag{5}$$

$C = (c^i_{i'})$ étant la matrice de passage.

D'ailleurs, la formule (5) est fournie par un calcul direct élémentaire: comme $\mathbf{e}_{i'} = c^i_{i'}\mathbf{e}_i$ et $\mathbf{e}_j = c^{j'}_j \mathbf{e}_{j'}$, on a $a^{j'}_{i'}\mathbf{e}_{j'} = \mathbf{A}\mathbf{e}_{i'} = c^i_{i'}\mathbf{A}\mathbf{e}_i = c^i_{i'}a^j_i\mathbf{e}_j = c^i_{i'}a^j_i c^{j'}_j \mathbf{e}_{j'}$, ce qui équivaut à (5). On n'a en fait que répéter les calculs de la leçon 5.

On passe aux notations matricielles en introduisant les vecteurs lignes

$$\mathbf{e} = (\mathbf{e}_1, \ldots, \mathbf{e}_n), \qquad \mathbf{e}' = (\mathbf{e}_{1'}, \ldots, \mathbf{e}_{n'}),$$
$$\mathbf{A}\mathbf{e} = (\mathbf{A}\mathbf{e}_1, \ldots, \mathbf{A}\mathbf{e}_n), \quad \mathbf{A}\mathbf{e}' = (\mathbf{A}\mathbf{e}_{1'}, \ldots, \mathbf{A}\mathbf{e}_{n'}),$$

auquel cas (cf. I.10, formule (14))

$$\mathbf{e}' = \mathbf{e}C, \qquad \mathbf{e} = \mathbf{e}'C^{-1}$$

et

$$\mathbf{A}\mathbf{e} = \mathbf{e}A, \quad \mathbf{A}\mathbf{e}' = \mathbf{e}'A'.$$

D'autre part, la propriété « l'opérateur $\mathbf{A}$ est linéaire » entraîne

$$\mathbf{A}\mathbf{e}' = \mathbf{A}(\mathbf{e}C) = (\mathbf{A}\mathbf{e})C.$$

Donc,

$$\mathbf{e}'A' = \mathbf{A}\mathbf{e}' = (\mathbf{A}\mathbf{e})C = \mathbf{e}AC = \mathbf{e}'C^{-1}AC,$$

et, partant, $A' = C^{-1}AC$. □

* * *

On dit qu'un opérateur $\mathbf{A}$ est *non dégénéré* si $\det A \neq 0$ (il est *dégénéré* dans le cas contraire). Le caractère correct de la définition découle de suite de la formule (5).

Les *opérateurs inversibles* sont d'un intérêt capital. Ce sont les opérateurs qui admettent un *opérateur inverse* $\mathbf{A}^{-1}$ tel que

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{E}.$$

Un opérateur $\mathbf{A}$ est dit *inversible à gauche* ou *inversible à droite* selon qu'il existe un opérateur $\mathbf{B}$ tel que

$$\mathbf{B}\mathbf{A} = \mathbf{E}$$

ou un opérateur $\mathbf{C}$ tel que

$$\mathbf{A}\mathbf{C} = \mathbf{E}.$$

Certains éléments inversibles des anneaux (ou des algèbres) arbitraires le sont soit à droite, soit à gauche, tandis qu'un opérateur linéaire est inversible s'il l'est au moins d'un côté. Cette propriété des opérateurs linéaires est étroitement liée au fait (qui est vraiment

curieux) qu'un opérateur linéaire est bijectif dès qu'il est injectif ou surjectif. (On observe à ce propos que si un opérateur inversible est évidemment bijectif, la propriété d'inversibilité de tout opérateur linéaire bijectif, i.e. la linéarité de l'opérateur inverse, ne tombe pas sous le sens.)

Proposition 2. *Etant donné un opérateur linéaire* $\mathbf{A} : \mathcal{V} \to \mathcal{V}$ *quelconque, on a les affirmations équivalentes*:

1° $\mathbf{A}$ *est inversible à gauche.*

2° $\mathbf{A}$ *est injectif, i.e.* $\operatorname{Ker} \mathbf{A} = 0$.

3° $\mathbf{A}$ *est inversible à droite.*

4° $\mathbf{A}$ *est surjectif, i.e.* $\operatorname{Im} \mathbf{A} = \mathcal{V}$.

5° $\mathbf{A}$ *est inversible.*

6° $\mathbf{A}$ *est bijectif.*

7° $\mathbf{A}$ *est non dégénéré.*

8° *Quelle que soit la base* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *de l'espace* $\mathcal{V}$, *les vecteurs* $\mathbf{Ae}_1, \ldots, \mathbf{Ae}_n$ *constituent une base.*

Démonstration. L'équivalence de 7° et 8° résulte de suite du théorème du rang d'une matrice. Quant aux autres affirmations, on procède selon le diagramme

$$\begin{tikzcd}
 & 1^\circ \arrow[r, Rightarrow] & 2^\circ \arrow[dl, Rightarrow] \\
5^\circ \arrow[ur, Rightarrow] \arrow[dr, Rightarrow] & & 8^\circ \arrow[ll, Rightarrow] \\
 & 3^\circ \arrow[r, Rightarrow] & 4^\circ \arrow[ul, Rightarrow]
\end{tikzcd}$$

$5^\circ \Rightarrow 1^\circ$. Si $\mathbf{A}^{-1}$ est l'inverse de $\mathbf{A}$, alors $\mathbf{A}^{-1}\mathbf{A} = \mathbf{E}$.

$1^\circ \Rightarrow 2^\circ$. Si $\mathbf{BA} = \mathbf{E}$ et $\mathbf{Ax} = \mathbf{0}$, alors $\mathbf{x} = \mathbf{Ex} = \mathbf{BAx} = \mathbf{B0} = \mathbf{0}$.

$2^\circ \Rightarrow 8^\circ$. S'il y a dépendance linéaire des vecteurs $\mathbf{Ae}_1, \ldots, \mathbf{Ae}_n$, i.e. si $k_1\mathbf{Ae}_1 + \ldots + k_n\mathbf{Ae}_n = 0$, où $(k_1, \ldots, k_n) \neq (0, \ldots, 0)$, le vecteur $\mathbf{e} = k_1\mathbf{e}_1 + \ldots + k_n\mathbf{e}_n \neq \mathbf{0}$ vérifie l'égalité $\mathbf{Ae} = \mathbf{0}$. Si $\operatorname{Ker} \mathbf{A} = \mathbf{0}$, les vecteurs $\mathbf{Ae}_1, \ldots, \mathbf{Ae}_n$ sont donc linéairement indépendants et forment une base.

$5^\circ \Rightarrow 3^\circ$. Si $\mathbf{A}^{-1}$ est l'inverse de $\mathbf{A}$, alors $\mathbf{AA}^{-1} = \mathbf{E}$.

$3^\circ \Rightarrow 4^\circ$. Si $\mathbf{AC} = \mathbf{E}$, alors $\mathbf{Ay} = \mathbf{x}$ pour tout vecteur $\mathbf{x} \in \mathcal{V}$, avec $\mathbf{y} = \mathbf{Cx}$.

$4^\circ \Rightarrow 8^\circ$. S'il existe pour tout $\mathbf{x} \in \mathcal{V}$ un vecteur $\mathbf{y} \in \mathcal{V}$ tel que $\mathbf{Ay} = \mathbf{x}$, alors $\mathbf{x} = y^1\mathbf{Ae}_1 + \ldots + y^n\mathbf{Ae}_n$, i.e. la famille de n vecteurs $\mathbf{Ae}_1, \ldots, \mathbf{Ae}_n$ est génératrice et constitue donc une base.

$8^\circ \Rightarrow 5^\circ$. Dans la base $\mathbf{e}'_1 = \mathbf{Ae}_1, \ldots, \mathbf{e}'_n = \mathbf{Ae}_n$, la famille $\mathbf{b}_1 = \mathbf{e}_1, \ldots, \mathbf{b}_n = \mathbf{e}_n$ définit un opérateur $\mathbf{B}$ pour lequel $\mathbf{Be}'_1 = \mathbf{b}_1, \ldots, \mathbf{Be}'_n = \mathbf{b}_n$, donc $(\mathbf{BA})\,\mathbf{e}_1 = \mathbf{e}_1, \ldots, (\mathbf{BA})\,\mathbf{e}_n = \mathbf{e}_n$, i.e. $\mathbf{BA} = \mathbf{E}$. On a également pour $\mathbf{B}$: $(\mathbf{AB})\,\mathbf{e}'_1 = \mathbf{e}'_1, \ldots, (\mathbf{AB})\,\mathbf{e}'_n = \mathbf{e}'_n$ et, partant, $\mathbf{AB} = \mathbf{E}$. Par conséquent, $\mathbf{A}$ est inversible (et $\mathbf{B} = \mathbf{A}^{-1}$). □

L'équation vectorielle

$$\mathbf{Ax} = \mathbf{b}$$

est représentée en coordonnées par un système de n équations linéaires à n inconnues. L'équivalence de 2° et 4° signifie donc en termes d'équations qu'*un système de n équations linéaires non homogènes à n inconnues est compatible pour les termes libres quelconques si et seulement si le système linéaire homogène associé n'admet pas d'autre solution à part la solution triviale.*

Des exemples simples montrent l'impossibilité de généraliser de suite ce beau résultat aux systèmes de n équations à m inconnues, $n \neq m$. On n'y arrive qu'après avoir reformulé l'affirmation.

* * *

Soit $\mathbf{A} \in \mathrm{Op}(\mathscr{V})$. On associe à un covecteur $\xi \in \mathscr{V}'$ quelconque une fonctionnelle $\mathbf{A}'\xi$ sur $\mathscr{V}$ si l'on pose

$$(\mathbf{A}'\xi)(\mathbf{x}) = \xi(\mathbf{A}\mathbf{x}), \quad \mathbf{x} \in \mathscr{V}. \tag{6}$$

On vérifie automatiquement que

a) la fonctionnelle $\mathbf{A}'\xi$ est linéaire, i.e. c'est un covecteur de $\mathscr{V}'$;

b) l'application intervenue $\mathbf{A}' : \mathscr{V}' \to \mathscr{V}'$ est linéaire, i.e. $\mathbf{A}'$ est un opérateur linéaire.

Définition 3. L'opérateur $\mathbf{A}'$ est dit *adjoint* de $\mathbf{A}$.

Si l'on introduit le couplage naturel $\langle \mathbf{x}, \xi \rangle = \xi(\mathbf{x})$ de $\mathscr{V}$ et $\mathscr{V}'$ (voir leçon 4), alors la formule (6) définissant $\mathbf{A}'$ s'écrit

$$\langle \mathbf{x}, \mathbf{A}'\xi \rangle = \langle \mathbf{A}\mathbf{x}, \xi \rangle.$$

La formule étant symétrique entraîne de suite que l'application $\mathbf{A} \mapsto \mathbf{A}'$ de l'espace $\mathrm{Op}(\mathscr{V})$ dans l'espace $\mathrm{Op}(\mathscr{V}')$ est *involutive*, i.e.

$$\mathbf{A}'' = \mathbf{A}.$$

Il en découle en particulier le caractère bijectif de $\mathbf{A} \mapsto \mathbf{A}'$. Il y a plus. Il est clair que

$$(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}' \text{ et } (k\mathbf{A})' = k\mathbf{A}'.$$

Cela signifie que *l'application $\mathbf{A} \mapsto \mathbf{A}'$ est un isomorphisme de l'espace vectoriel* $\mathrm{Op}(\mathscr{V})$ *sur l'espace vectoriel* $\mathrm{Op}(\mathscr{V}')$. □

Ainsi, $\mathrm{Op}(\mathscr{V})$ et $\mathrm{Op}(\mathscr{V}')$ sont reliés par un isomorphisme canonique contrairement aux espaces $\mathscr{V}$ et $\mathscr{V}'$!

L'application $\mathbf{A} \mapsto \mathbf{A}'$ n'est pas un isomorphisme pour la multiplication car elle change l'ordre des facteurs:

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'.$$

En effet, $\langle \mathbf{x}, (\mathbf{AB})'\xi \rangle = \langle \mathbf{ABx}, \xi \rangle = \langle \mathbf{Bx}, \mathbf{A}'\xi \rangle = \langle \mathbf{x}, \mathbf{B}'\mathbf{A}'\xi \rangle$. Un isomorphisme linéaire d'algèbres muni de cette propriété s'appelle d'ordinaire un *anti-isomorphisme*.

La formule $a_i^j = \mathbf{e}^j(\mathbf{A}\mathbf{e}_i)$ pour les éléments de la matrice de $\mathbf{A}$ dans la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ signifie que

$$a_i^j = \langle \mathbf{A}\mathbf{e}_i, \mathbf{e}^j \rangle,$$

si bien qu'on a pour les éléments $a_i'^j$ de la matrice de $\mathbf{A}'$ dans la base duale $\mathbf{e}^1, \ldots, \mathbf{e}^n$:

$$a_i'^j = \langle \mathbf{e}_i, \mathbf{A}'\mathbf{e}^j \rangle = \langle \mathbf{A}\mathbf{e}_i, \mathbf{e}^j \rangle,$$

donc $a_i'^j = a_i^j$, i.e. $\mathbf{A}\mathbf{e}^j = a_i^j \mathbf{e}^i$. Cela *ne veut pas dire* que les opérateurs $\mathbf{A}$ et $\mathbf{A}'$ ont même matrice. En effet, les colonnes de la matrice d'un opérateur sont par définition les coordonnées des vecteurs résultant de l'action de l'opérateur sur les vecteurs de base. S'agissant de $\mathbf{A}$, cela signifie (en vertu de $\mathbf{A}\mathbf{e}_i = a_i^j \mathbf{e}_j$) que la colonne i de sa matrice est formée de nombres $a_i^1, \ldots, a_i^n$. Quant à l'opérateur $\mathbf{A}'$, la formule $\mathbf{A}'\mathbf{e}^j = a_i^j \mathbf{e}^i$ équivaut à dire que la colonne j de la matrice de $\mathbf{A}'$ a pour éléments les nombres $a_1^j, \ldots, a_n^j$, i.e. les nombres qui constituent la *ligne* j de la matrice de $\mathbf{A}$. Ainsi, la matrice de $\mathbf{A}'$ rapportée à la base duale $\mathbf{e}^1, \ldots, \mathbf{e}^n$ est la *transposée* A^{T} de la matrice A de $\mathbf{A}$ dans la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

Proposition 3.

$$\operatorname{Ker} \mathbf{A}' = (\operatorname{Im} \mathbf{A})^\circ, \qquad \operatorname{Im} \mathbf{A}' = (\operatorname{Ker} \mathbf{A})^\circ,$$
$$\operatorname{Ker} \mathbf{A} = (\operatorname{Im} \mathbf{A}')^\circ, \qquad \operatorname{Im} \mathbf{A} = (\operatorname{Ker} \mathbf{A}')^\circ.$$

Démonstration. Dire que $\xi \in \operatorname{Ker} \mathbf{A}'$ c'est dire que tout vecteur $\mathbf{x} \in \mathcal{V}$ satisfait à l'égalité $(\mathbf{A}'\xi)(\mathbf{x}) = 0$, i.e. à $\xi(\mathbf{A}\mathbf{x}) = 0$ caractérisant les covecteurs de $(\operatorname{Im} \mathbf{A})^\circ$. Par conséquent, $\operatorname{Ker} \mathbf{A}' = (\operatorname{Im} \mathbf{A})^\circ$. On remplace $\mathbf{A}$ par $\mathbf{A}'$, ce qui donne $\operatorname{Ker} \mathbf{A} = (\operatorname{Im} \mathbf{A}')^\circ$. On passe aux annulateurs (et on utilise la proposition 5 de la leçon 4), il vient $(\operatorname{Ker} \mathbf{A}')^\circ = \operatorname{Im} \mathbf{A}$ et $(\operatorname{Ker} \mathbf{A})^\circ = \operatorname{Im} \mathbf{A}'$. □

* * *

En particulier, $\operatorname{Im} \mathbf{A} = \mathcal{V}$ si et seulement si $\operatorname{Ker} \mathbf{A}' = 0$. En passant aux coordonnées, ce résultat s'énonce (pour $m = n$) comme suit : *le système d'équations linéaires non homogènes*

$$(7) \qquad \begin{array}{c} a_1^1 x_1 + \ldots + a_1^n x_n = b_1, \\ \cdots\cdots\cdots\cdots\cdots \\ a_m^1 x_1 + \ldots + a_m^n x_n = b_m \end{array}$$

est compatible pour tous les $b_1, \ldots, b_m$ *si et seulement si le système d'équations homogènes*

$$(8) \qquad \begin{array}{c} a_1^1 x_1 + \ldots + a_m^1 x_m = 0, \\ \cdots\cdots\cdots\cdots\cdots \\ a_1^n x_1 + \ldots + a_m^n x_m = 0 \end{array}$$

de matrice transposée n'a pas d'autre solution que la solution triviale. □

Cette affirmation connue sous le nom d'*alternative de Fredholm* est juste (on le constate sans peine) pour n'importe quels m et n.

En effet, le système (8) n'admet pas d'autre solution que la solution triviale si et seulement si la matrice de ses coefficients est de rang m ($r = m$). D'autre part, le théorème de Kronecker-Capelli établit que le système (7) est compatible sous la condition nécessaire et suffisante que le rang r de la matrice des coefficients soit conservé par adjonction de la colonne $\mathbf{b}$ des seconds membres, ce qui a évidemment pour toute colonne $\mathbf{b}$ si et seulement si $r = m$. □

Afin d'énoncer l'alternative de Fredholm en termes d'opérateurs pour $m \neq n$, on généralise la notion d'opérateur adjoint à un homomorphisme $\varphi : \mathscr{V} \to \mathscr{W}$ quelconque, où $\mathscr{W} \neq \mathscr{V}$. La généralisation s'effectue sans difficulté au prix de modifications mineures. Cela étant, l'analogue de la proposition 3 reste en vigueur, ce qui fournit l'alternative de Fredholm sous forme générale.

* * *

Définition 4. Un sous-espace $\mathscr{P}$ de l'espace $\mathscr{V}$ est dit *invariant* par un opérateur $\mathbf{A} : \mathscr{V} \to \mathscr{V}$ si

$$\mathbf{Ax} \in \mathscr{P} \text{ pour tout vecteur } \mathbf{x} \in \mathscr{P}.$$

On définit l'opérateur

$$\mathbf{A}\,|_{\mathscr{P}} \in \mathrm{Op}\,(\mathscr{P})$$

agissant par la formule

$$(\mathbf{A}|_{\mathscr{P}})\,\mathbf{x} = \mathbf{Ax}, \quad \mathbf{x} \in \mathscr{P},$$

le vecteur $\mathbf{Ax}$ étant considéré comme élément de $\mathscr{P}$.

L'opérateur $\mathbf{A}|_{\mathscr{P}}$ s'appelle la *restriction* de l'opérateur $\mathbf{A}$ au sous-espace invariant $\mathscr{P}$. On dit également que $\mathbf{A}|_{\mathscr{P}}$ est *induit* par $\mathbf{A}$.

Comme $\dim \mathscr{P} < \dim \mathscr{V}$, l'opérateur $\mathbf{A}|_{\mathscr{P}}$ se prête mieux à l'étude que $\mathbf{A}$. Par ailleurs, l'examen de $\mathbf{A}|_{\mathscr{P}}$ apporte beaucoup de renseignements sur $\mathbf{A}$.

La situation est particulièrement favorable avec (s'il existe) un deuxième sous-espace invariant $\mathscr{Q}$ supplémentaire de $\mathscr{P}$, i.e. avec $\mathscr{V}$ qui est somme directe $\mathscr{V} = \mathscr{P} \oplus \mathscr{Q}$ de sous-espaces invariants $\mathscr{P}$ et $\mathscr{Q}$. Dans ce cas $\mathbf{A}$ est bien reconstitué à partir de $\mathbf{A}\,|_{\mathscr{P}}$ et $\mathbf{A}\,|_{\mathscr{Q}}$. En effet, on a évidemment pour tout vecteur $\mathbf{z} = \mathbf{x} + \mathbf{y}$ de $\mathscr{V}$, où $\mathbf{x} \in \mathscr{P}$, $\mathbf{y} \in \mathscr{Q}$,

$$\mathbf{Az} = (\mathbf{A}\,|_{\mathscr{P}})\,\mathbf{x} + (\mathbf{A}\,|_{\mathscr{Q}})\,\mathbf{y}.$$

Le fait que $\mathbf{A}$ se réduit à ses restrictions $\mathbf{A}\,|_{\mathscr{P}}$ et $\mathbf{A}\,|_{\mathscr{Q}}$ est mis en évidence par l'examen de la matrice $A = (a_i^j)$ de $\mathbf{A}$ dans une base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de $\mathscr{V}$ telle que $\mathscr{P} = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$ et $\mathscr{Q} = [\mathbf{e}_{p+1}, \ldots, \mathbf{e}_n]$. En effet,

$$a_i^j = 0 \text{ si } 1 \leqslant i \leqslant p \text{ et } p+1 \leqslant j \leqslant n$$

parce que $\mathbf{A}\mathbf{e}_i = a_i^j \mathbf{e}_j \in \mathcal{P}$ pour $1 \leqslant i \leqslant p$. De même,

$$a_i^j = 0 \text{ si } p+1 \leqslant i \leqslant n \text{ et } 1 \leqslant j \leqslant p$$

du moment que $\mathbf{A}\mathbf{e}_i \in \mathcal{Q}$ pour $p+1 \leqslant i \leqslant n$. Cela signifie que la matrice A rapportée à la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ a la *forme bloc-diagonale*

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \tag{9}$$

avec A_1 la matrice de $\mathbf{A}\,|\,_{\mathcal{P}}$ dans $\mathbf{e}_1, \ldots, \mathbf{e}_p$ et A_2 la matrice de $\mathbf{A}\,|\,_{\mathcal{Q}}$ rapportée à la base $\mathbf{e}_{p+1}, \ldots, \mathbf{e}_n$.

On dit des fois d'une matrice de la forme (9) qu'*elle est décomposée en somme directe des matrices A_1 et A_2* (ce qui se note $A = A_1 \oplus A_2$). Ainsi, chaque décomposition de $\mathcal{V}$ en somme directe de sous-espaces invariants définit la décomposition de la matrice d'un opérateur en somme directe des opérateurs induits.

Si le sous-espace invariant $\mathcal{P}$ n'admet pas de supplémentaire invariant $\mathcal{Q}$ (ou si on ne sait rien sur $\mathcal{Q}$), on met la matrice A (en prenant la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de sorte que $\mathcal{P} = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$) *sous forme bloc-triangulaire*

$$A = \begin{pmatrix} A_1 & C \\ 0 & B \end{pmatrix}, \tag{10}$$

A_1 étant la matrice de l'opérateur $\mathbf{A}\,|\,_{\mathcal{P}}$.

L'invariance de $\mathcal{P}$ par $\mathbf{A}$ entraîne de suite que la formule

$$\mathbf{B}\,(\mathbf{x} + \mathcal{P}) = \mathbf{A}\mathbf{x} + \mathcal{P}$$

définit bien sur l'espace quotient $\mathcal{V}/\mathcal{P}$ un opérateur (manifestement linéaire)

$$\mathbf{B} : \mathcal{V}/\mathcal{P} \to \mathcal{V}/\mathcal{P}$$

dont on dit également qu'il est *induit* par $\mathbf{A}$.

Si la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de l'espace $\mathcal{V}$ est choisie de façon que $\mathcal{P} = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$, alors les classes $\mathbf{e}_{p+1} + \mathcal{P}, \ldots, \mathbf{e}_n + \mathcal{P}$ de $\mathcal{V}$ suivant $\mathcal{P}$ constituent évidemment une base de l'espace quotient $\mathcal{V}/\mathcal{P}$, et la matrice de l'opérateur $\mathbf{B}$ dans cette base est la matrice B de (10).

LEÇON 15

Valeurs propres. — Racines caractéristiques. — Opérateurs diagonalisables. — Opérateurs à spectre simple. — Existence d'une base dans laquelle la matrice d'un opérateur est triangulaire. — Opérateurs nilpotents.

Les sous-espaces invariants les plus simples sont des sous-espaces de dimension 1.

Définition 1. Un vecteur $\mathbf{x} \neq \mathbf{0}$ s'appelle *vecteur propre* d'un opérateur $\mathbf{A}$ s'il engendre un sous-espace invariant de dimension 1.

La condition nécessaire et suffisante en est évidemment l'existence d'un nombre $\lambda \in \mathbb{K}$ tel que

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}. \tag{1}$$

On appelle *valeur propre* d'un opérateur $\mathbf{A}$ tout nombre $\lambda \in \mathbb{K}$ tel qu'il existe un vecteur $\mathbf{x} \neq \mathbf{0}$ vérifiant (1) (et qui est donc un vecteur propre de $\mathbf{A}$). On dit aussi que $\mathbf{x}$ est un vecteur propre *associé à* la valeur propre λ.

Il y a intérêt à considérer le vecteur nul (qui n'est pas par définition un vecteur propre) comme relatif à chaque valeur propre. Dans ce cas, l'ensemble $\mathcal{P}_\lambda$ de tous les vecteurs $\mathbf{x} \in \mathcal{V}$ correspondant à une valeur propre λ quelconque constitue évidemment un sous-espace appelé *sous-espace propre associé à la valeur propre* λ. La dimension $p_\lambda = \dim \mathcal{P}_\lambda$ du sous-espace propre s'appelle la *multiplicité géométrique* de la valeur propre λ. Par définition, on a $1 \leqslant p_\lambda \leqslant n$.

Le sous-espace invariant $[\mathbf{x}]$ de dimension 1 engendré par tout vecteur propre $\mathbf{x} \neq \mathbf{0}$ associé à une valeur propre λ est tout entier dans $\mathcal{P}_\lambda$. Inversement, chaque sous-espace de dimension 1 de $\mathcal{P}_\lambda$ est invariant, si bien que l'*espace* $\mathcal{P}_\lambda$ *est* en particulier *somme directe de sous-espaces invariants unidimensionnels*. Il suffit, pour obtenir cette décomposition, de choisir dans $\mathcal{P}_\lambda$ une base arbitraire.

Géométriquement, $\mathcal{P}_\lambda$ est le plus grand sous-espace invariant sur lequel l'opérateur $\mathbf{A}$ (plus précisément, sa restriction $\mathbf{A}|_{\mathcal{P}_\lambda}$) est un opérateur scalaire $\lambda\mathbf{E}$. On dit également que $\mathcal{P}_\lambda$ *est le noyau de l'opérateur* $\mathbf{A} - \lambda\mathbf{E}$:

$$\mathcal{P}_\lambda = \operatorname{Ker}(\mathbf{A} - \lambda\mathbf{E}).$$

En effet, l'égalité $(\mathbf{A} - \lambda\mathbf{E})\,\mathbf{x} = \mathbf{0}$ est exactement l'égalité (1). □

* * *

Ainsi, un nombre $\lambda \in \mathbb{K}$ est valeur propre de A si et seulement si l'opérateur $\mathbf{A} - \lambda\mathbf{E}$ a son noyau non nul, i.e. s'il est non inversible (dégénéré) ; voir leçon 14, prop. 2. Autrement dit, λ est valeur propre si et seulement si

$$\det (A - \lambda E) = 0,$$

A étant la matrice de l'opérateur A rapportée à une base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ quelconque.

Le déterminant

$$\det (A - \lambda E) = \begin{vmatrix} a_1^1 - \lambda & \ldots & a_n^1 \\ \cdot & \cdot & \cdot \\ a_1^n & \ldots & a_n^n - \lambda \end{vmatrix}$$

est, on le voit aisément, un polynôme de degré n en λ. *Ce polynôme est indépendant de la base* $\mathbf{e}_1, \ldots, \mathbf{e}_n$. En effet, la matrice de l'opérateur A dans toute autre base s'écrit $C^{-1}AC$ (voir formule (5), leçon 14) et

$$C^{-1}AC - \lambda E = C^{-1}(A - \lambda E)\, C,$$

si bien que

$$\det (C^{-1}AC - \lambda E) = (\det C)^{-1} \det (A - \lambda E) (\det C)^{-1} = \\ = \det (A - \lambda E). \ \square$$

Définition 2. Le polynôme

$$f_{\mathbf{A}}(\lambda) = \det (A - \lambda E)$$

s'appelle le *polynôme caractéristique* de l'opérateur A, et ses racines (dans l'extension correspondante de $\mathbb{K}$) en constituent les *racines caractéristiques*.

Les résultats ci-dessus autorisent à dire que *toute valeur propre de l'opérateur* A *en est une racine caractéristique*, et, inversement, *toute racine caractéristique dans* $\mathbb{K}$ *est une valeur propre.* $\square$

Cette affirmation (et le fait que $\mathcal{P}_\lambda = \operatorname{Ker}(\mathbf{A} - \lambda\mathbf{E})$) inspire un procédé de recherche des sous-espaces propres. On commence par résoudre l'équation $f_{\mathbf{A}}(\lambda) = 0$, ce qui donne toutes ses racines dans $\mathbb{K}$. A l'étape suivante, on trouve pour chaque racine λ_i le sous-espace $\mathcal{P}_{\lambda_i}$ en résolvant un système d'équations linéaires homogènes de matrice $A - \lambda_i E$.

La multiplicité de la valeur propre λ_0 en tant que racine du polynôme caractéristique, i.e. un nombre n_{λ_0} tel que le polynôme $f_{\mathbf{A}}(\lambda)$ soit divisible par $(\lambda - \lambda_0)^{n_{\lambda_0}}$, mais qu'il ne le soit pas par $(\lambda - \lambda_0)^{n_{\lambda_0}+1}$, s'appelle la *multiplicité algébrique* de la valeur propre λ_0. On vérifie de suite que *la multiplicité algébrique d'une valeur propre est au moins égale à sa multiplicité géométrique* :

$$p_{\lambda_0} \leqslant n_{\lambda_0}.$$

En effet, soit $p = p_{\lambda_0}$ et soit $\mathbf{e}_1, \ldots, \mathbf{e}_n$ une base de l'espace $\mathscr{V}$ telle que $\mathscr{P}_{\lambda_0} = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$. La matrice de $\mathbf{A}$ dans cette base s'écrit

$$\begin{pmatrix} A_1 & C \\ 0 & B \end{pmatrix}, \tag{2}$$

si bien que

$$f_{\mathbf{A}}(\lambda) = \det(A - \lambda E) = \det(A_1 - \lambda E) \cdot \det(B - \lambda E).$$

Or A_1 est la matrice de

$$\mathbf{A}|_{\mathscr{P}_{\lambda_0}} = \lambda_0 \mathbf{E},$$

donc $\det(A_1 - \lambda E) = (\lambda_0 - \lambda)^p$, ce qui démontre la divisibilité de $f_{\mathbf{A}}(\lambda)$ par $(\lambda - \lambda_0)^p$ et, partant, l'inégalité $p \leqslant n_{\lambda_0}$. □

Remarque. L'opérateur $\mathbf{A}$ est représenté par une matrice de la forme (2) dans toute base pour laquelle $\mathscr{P} = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$ est invariant. Cela étant, A_1 est la matrice de $\mathbf{A}|_{\mathscr{P}}$ et B la matrice de l'opérateur $\mathbf{B}: \mathscr{V}/\mathscr{P} \to \mathscr{V}/\mathscr{P}$ induit. Ainsi, *on a pour tout sous-espace invariant* $\mathscr{P} \subset \mathscr{V}$

$$f_{\mathbf{A}}(\lambda) = f_{\mathbf{A}|_{\mathscr{P}}}(\lambda)\, f_{\mathbf{B}}(\lambda).$$

Soient $\lambda_1, \ldots, \lambda_m$ des valeurs propres *distinctes* de l'opérateur $\mathbf{A}$ et

$$\mathscr{P}_1 = \mathscr{P}_{\lambda_1}, \ldots, \mathscr{P}_m = \mathscr{P}_{\lambda_m}$$

les sous-espaces propres associés.

Proposition 1. *La somme*

$$\mathscr{P} = \mathscr{P}_1 + \ldots + \mathscr{P}_m$$

des sous-espaces $\mathscr{P}_1, \ldots, \mathscr{P}_m$ *est directe, i.e. l'égalité*

$$\mathbf{x}_1 + \ldots + \mathbf{x}_m = \mathbf{0}, \tag{3}$$

avec $\mathbf{x}_1 \in \mathscr{P}_1, \ldots, \mathbf{x}_m \in \mathscr{P}_m$, *a lieu si et seulement si*

$$\mathbf{x}_1 = \mathbf{0}, \ldots, \mathbf{x}_m = \mathbf{0}.$$

Démonstration. On raisonne par récurrence sur m. L'affirmation est évidente pour $m = 1$ (elle est d'ailleurs vide de signification). Supposons-la vraie pour la somme de $m - 1$ espaces $\mathscr{P}_1, \ldots, \mathscr{P}_{m-1}$. On applique $\mathbf{A}$ à l'égalité (3):

$$\lambda_1 \mathbf{x}_1 + \ldots + \lambda_m \mathbf{x}_m = \mathbf{0}. \tag{4}$$

On en retranche (3) multiplié par λ_m:

$$(\lambda_1 - \lambda_m)\mathbf{x}_1 + \ldots + (\lambda_{m-1} - \lambda_m)\mathbf{x}_{m-1} = \mathbf{0}.$$

Il en résulte par hypothèse de récurrence

$$(\lambda_1 - \lambda_m)\, x_1 = 0, \ldots, (\lambda_{m-1} - \lambda_m)\, x_{m-1} = 0,$$

donc

$$x_1 = 0, \ldots, x_{m-1} = 0$$

(car $\lambda_1 - \lambda_m \neq 0, \ldots, \lambda_{m-1} - \lambda_m \neq 0$ par hypothèse), auquel cas $x_m = 0$ d'après (3). □

* * *

On suppose qu'il existe des valeurs propres (distinctes)

(5) $$\lambda_1, \ldots, \lambda_m$$

telles que

(6) $$\mathcal{P}_{\lambda_1} \oplus \ldots \oplus \mathcal{P}_{\lambda_m} = \mathcal{V},$$

et, partant,

(7) $$p_{\lambda_1} + \ldots + p_{\lambda_m} = n.$$

On établit facilement que *les nombres* (5) *sont les seules valeurs propres de l'opérateur* A. En effet, s'il existe de plus une valeur propre λ_0 distincte des valeurs propres données, le sous-espace $\mathcal{P}_{\lambda_0}$ et $\mathcal{V}$ forment, selon la proposition 1, une somme directe, ce qui est impossible. □

On choisit une base dans chaque $\mathcal{P}_{\lambda_1}, \ldots, \mathcal{P}_{\lambda_m}$ et on obtient une base de $\mathcal{V}$ dont les éléments sont vecteurs propres. La matrice de A rapportée à cette base est diagonale:

(8) $$\begin{pmatrix} \lambda_1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & \lambda_m \end{pmatrix},$$

et elle a pour éléments diagonaux les valeurs propres (5), chaque λ_i étant comptée p_{λ_i} fois.

Inversement, soit une base de $\mathcal{V}$ telle que la matrice A de l'opérateur A rapportée à la base soit diagonale. Les vecteurs de base sont dans ce cas vecteurs propres, et les éléments diagonaux de A constituent des valeurs propres de A. Soient $\lambda_1, \ldots, \lambda_m$ les éléments diagonaux distincts de A, l'élément λ_i, $i = 1, \ldots, m$, étant compté q_i fois. Soit ensuite $\mathcal{Q}_i$, $i = 1, \ldots, m$, le sous-espace de $\mathcal{V}$ engendré par les vecteurs de base associés à la valeur propre λ_i. Alors $\dim \mathcal{Q}_i = q_i$,

$$\mathcal{Q}_1 \oplus \ldots \oplus \mathcal{Q}_m = \mathcal{V}$$

et $\mathcal{G}_i \subset \mathcal{P}_{\lambda_i}$. Aussi on a, en particulier,

$$q_1 + \ldots + q_m = n \text{ et } q_1 \leqslant p_{\lambda_1}, \ldots, q_m \leqslant p_{\lambda_m}. \tag{9}$$

Mais la somme $\mathcal{P}_{\lambda_1} + \ldots + \mathcal{P}_{\lambda_m}$ des sous-espaces $\mathcal{P}_{\lambda_1}, \ldots, \mathcal{P}_{\lambda_m}$ est par la proposition 1 une somme directe, d'où sa dimension $p_{\lambda_1} + \ldots + p_{\lambda_m}$. Ainsi, $p_{\lambda_1} + \ldots + p_{\lambda_m} \leqslant n$, ce qui entraîne par suite de (9)

$$q_1 = p_{\lambda_1}, \ldots, q_m = p_{\lambda_m},$$

i.e.

$$\mathcal{G}_1 = \mathcal{P}_{\lambda_1}, \ldots, \mathcal{G}_m = \mathcal{P}_{\lambda_m}.$$

Par conséquent, on a pour $\mathcal{P}_{\lambda_1}, \ldots, \mathcal{P}_{\lambda_m}$ la décomposition (6).

Comme l'existence d'une base dans laquelle la matrice de l'opérateur A est diagonale est équivalente à la propriété de $\mathcal{V}$ de se décomposer en somme directe de sous-espaces invariants de dimension 1, on vient de démontrer la

Proposition 2. *Etant donné un opérateur linéaire* A : $\mathcal{V} \to \mathcal{V}$ *quelconque, les affirmations suivantes sont équivalentes*:

1° *Il existe des valeurs propres* $\lambda_1, \ldots, \lambda_m$ *telles que*

$$\mathcal{P}_{\lambda_1} \oplus \ldots \oplus \mathcal{P}_{\lambda_m} = \mathcal{V}.$$

2° *L'espace* $\mathcal{V}$ *est somme directe de sous-espaces de dimension* 1 *invariants par l'opérateur* A.

3° *On trouve dans* $\mathcal{V}$ *une base de vecteurs propres, i.e. une base dans laquelle la matrice de* A *est diagonale.* □

Cela étant, l'opérateur A n'a pas d'autres valeurs propres que les valeurs propres $\lambda_1, \ldots, \lambda_m$ intervenant dans 1° (et de façon indirecte dans 2°). Chaque base par rapport à laquelle la matrice de A est diagonale est réunion des bases des espaces $\mathcal{P}_{\lambda_1}, \ldots, \mathcal{P}_{\lambda_m}$, si bien qu'il existe dans cette base, pour toute valeur propre λ_i, exactement p_{λ_i} vecteurs associés à λ_i.

Définition 3. On dit qu'un opérateur A est *diagonalisable* si l'on a pour lui les affirmations 1° à 3° de la proposition 2.

On obtient de suite pour le polynôme caractéristique d'un opérateur diagonalisable A dans la base de vecteurs propres:

$$f_A(\lambda) = (\lambda - \lambda_1)^{p_1} \ldots (\lambda - \lambda_m)^{p_m},$$

avec $\lambda_1, \ldots, \lambda_m$ les valeurs propres de A et $p_1 = p_{\lambda_1}, \ldots, p_m = p_{\lambda_m}$ leurs multiplicités géométriques respectives. Ainsi, *toute racine caractéristique* λ_0 *d'un opérateur diagonalisable est dans le corps* K (i.e. c'est valeur propre), *et sa multiplicité algébrique* n_{λ_0} *est égale à sa multiplicité géométrique* p_{λ_0}. □

Il se trouve que cette condition nécessaire d'opérateurs diagonalisables est encore suffisante, si bien qu'on a le

Théorème 1. *Un opérateur linéaire* **A** *est diagonalisable si et seulement si toute racine caractéristique* λ_0 *de* **A** *est dans* $\mathbb{K}$ *et* $n_{\lambda_0} = p_{\lambda_0}$.

Démonstration. On démontre la suffisance seule.

Soient $\lambda_1, \ldots, \lambda_m$ toutes les racines caractéristiques de **A**. Elles sont par hypothèse dans $\mathbb{K}$ et constituent donc autant de valeurs propres. Par conséquent, on définit les sous-espaces $\mathcal{P}_{\lambda_1}, \ldots, \mathcal{P}_{\lambda_m}$ dont la somme (qui est, on le sait, somme directe) est de dimension

$$p_{\lambda_1} + \ldots + p_{\lambda_m} = n_{\lambda_1} + \ldots + n_{\lambda_m} = n$$

(la somme des multiplicités des racines d'un polynôme est égale au degré de celui-ci). Donc $\mathcal{P}_{\lambda_1} \oplus \ldots \oplus \mathcal{P}_{\lambda_m} = \mathcal{V}$. □

* * *

Définition 4. L'ensemble des racines caractéristiques d'un opérateur **A** s'appelle le *spectre* de **A**, qui est dit *simple* si chaque racine caractéristique λ_0 est racine simple du polynôme caractéristique, i.e. si $n_{\lambda_0} = 1$.

On dit que *le spectre est dans* $\mathbb{K}$ s'il en est ainsi de toutes les racines caractéristiques.

Proposition 3. *Tout opérateur à spectre simple dans* $\mathbb{K}$ *est diagonalisable.*

Démonstration. Comme $1 \leqslant p_\lambda \leqslant n_\lambda$, on a nécessairement $p_\lambda = 1$ pour $n_\lambda = 1$ (et donc $p_\lambda = n_\lambda$). □

Cette condition d'opérateurs diagonalisables n'est pas nécessaire, mais elle est très commode dans la pratique.

* * *

Soit $\mathcal{P} \subset \mathcal{V}$ quelconque invariant (par un opérateur **A**). Selon la remarque ci-dessus, le polynôme caractéristique $f_{\mathbf{B}}(\lambda)$ de l'opérateur induit $\mathbf{B} : \mathcal{V}/\mathcal{P} \to \mathcal{V}/\mathcal{P}$ divise le polynôme caractéristique $f_{\mathbf{A}}(\lambda)$ de l'opérateur **A**, si bien que chaque racine caractéristique de **B** est racine caractéristique de multiplicité algébrique au moins égale de **A**. Si **A** a en particulier son spectre dans $\mathbb{K}$, il en est de même du spectre de **B** qui possède donc au moins une valeur propre λ_0. Soit $\mathbf{x}_0 + \mathcal{P}$ un vecteur propre associé. L'égalité $\mathbf{B}(\mathbf{x}_0 + \mathcal{P}) = \lambda_0(\mathbf{x}_0 + \mathcal{P})$ signifie que $\mathbf{A}\mathbf{x}_0 = \lambda_0\mathbf{x}_0 + \mathbf{a}_0$, $\mathbf{a}_0 \in \mathcal{P}$, d'où l'invariance par **A** du sous-espace $\mathcal{Q}$ engendré par $\mathcal{P}$ et le vecteur $\mathbf{x}_0$ (i.e. formé de tous les vecteurs $k\mathbf{x}_0 + \mathbf{a}$, $k \in \mathbb{K}$, $\mathbf{a} \in \mathcal{P}$; on note que $\mathbf{x}_0 \notin \mathcal{P}$). Du moment que $\dim \mathcal{Q} = \dim \mathcal{P} + 1$, cela démontre la

Proposition 4. *Si le spectre d'un opérateur linéaire* $\mathbf{A} : \mathcal{V} \to \mathcal{V}$ *est dans* $\mathbb{K}$, *tout sous-espace invariant par* **A** *est contenu dans le sous-espace invariant de dimension supérieure d'une unité.* □

On construit donc, à partir de $\mathcal{P}_0 = 0$, la chaîne croissante de sous-espaces invariants

$$0 = \mathcal{P}_0 \subset \mathcal{P}_1 \subset \ldots \subset \mathcal{P}_n = \mathcal{V}$$

de dimension 0, 1, . . ., n respectivement. Il est clair que la matrice de l'opérateur A dans la base correspondante $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de $\mathcal{V}$, i.e. dans une base telle que $\mathcal{P}_i = [\mathbf{e}_1, \ldots, \mathbf{e}_i]$ pour tout $i = 1, \ldots, n$, est la matrice triangulaire

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix} \tag{10}$$

ayant pour éléments diagonaux les valeurs propres de A, chacune d'elles étant comptée avec sa multiplicité algébrique. Cela prouve la

Proposition 5. *Etant donné un opérateur linéaire* $\mathbf{A} : \mathcal{V} \to \mathcal{V}$ *quelconque à spectre dans* $\mathbb{K}$, *il existe dans* $\mathcal{V}$ *une base par rapport à laquelle la matrice de* A *est triangulaire.*

Si $\mathbb{K} = \mathbb{C}$, le résultat s'applique certes à tout opérateur linéaire.

* * *

On obtient un résultat plus précis pour les opérateurs d'une classe spéciale.

Définition 5. Un opérateur A (resp. une matrice A) est dit *nilpotent* (resp. *nilpotente*) s'il existe un entier naturel m tel que $\mathbf{A}^m = 0$ (resp. $A^m = 0$). Le plus petit m constitue l'*ordre de nilpotence* de l'opérateur (resp. matrice).

On constate facilement *la nullité de toute valeur propre d'un opérateur nilpotent*. En effet, si $\mathbf{Ax} = \lambda\mathbf{x}$, alors $\mathbf{A}^k\mathbf{x} = \lambda^k\mathbf{x}$ pour tout k, si bien que λ^m est nécessairement nul pour $\mathbf{A}^m = 0$ et $\mathbf{x} \neq 0$, i.e. $\lambda = 0$. □

Un opérateur nilpotent non nul n'est donc pas diagonalisable.

Un opérateur nilpotent est par exemple un opérateur tel qu'il existe un vecteur $\mathbf{e} = 0$ ayant la propriété suivante : les vecteurs

$$\mathbf{e}, \mathbf{Ae}, \ldots, \mathbf{A}^{n-1}\mathbf{e}$$

constituent une base de $\mathcal{V}$ et $\mathbf{A}^n\mathbf{e} = 0$. Dans la base

$$\mathbf{e}_1 = \mathbf{A}^{n-1}\mathbf{e}, \ldots, \mathbf{e}_{n-1} = \mathbf{Ae}, \ \mathbf{e}_n = \mathbf{e},$$

cet opérateur est représenté par la matrice

$$\begin{pmatrix} 0 & 1 & & 0 \\ & 0 & 1 & \\ & & \ddots & \\ & & & 1 \\ 0 & & & 0 \end{pmatrix}. \tag{11}$$

Ces opérateurs sont dits *cycliques*.

Quels que soient $\mathbf{x} = x^1\mathbf{e}_1 + \ldots + x^n\mathbf{e}_n$ et $m \leqslant n$, on a $\mathbf{A}^m\mathbf{x} = x^{m+1}\mathbf{e}_1 + \ldots + x^n\mathbf{e}_{n-m}$, et, en particulier, $\mathbf{A}^n\mathbf{x} = \mathbf{0}$. Ainsi, *un opérateur cyclique est nilpotent d'ordre de nilpotence n.* □

Si $n = 1$, les opérateurs cycliques sont nuls.

Il se trouve qu'un opérateur nilpotent quelconque se réduit à un opérateur cyclique.

Proposition 6. *Pour tout opérateur nilpotent* $\mathbf{A} : \mathcal{V} \to \mathcal{V}$, *l'espace* $\mathcal{V}$ *se décompose en somme directe de sous-espaces invariants*:

$$\mathcal{V} = \mathcal{P}_1 \oplus \ldots \oplus \mathcal{P}_m$$

sur chacun desquels $\mathbf{A}$ *induit un opérateur cyclique.*

La démonstration sera donnée dans la leçon suivante.

LEÇON 16

Décomposition d'un opérateur nilpotent en somme directe d'opérateurs cycliques. — Sous-espaces principaux. — Forme normale de Jordan. — Théorème de Hamilton-Cayley.

Soient $\mathbf{A}: \mathscr{V} \to \mathscr{V}$ un opérateur nilpotent arbitraire et

$$\mathscr{P}_i = \operatorname{Im} \mathbf{A}^i, \quad i = 0, 1, \ldots, m,$$

m étant l'ordre de nilpotence de $\mathbf{A}$. Comme $\mathbf{A}^{i+1}\mathbf{x} = \mathbf{A}^i(\mathbf{A}\mathbf{x})$, on a

$$\mathbf{0} = \mathscr{P}_m \subset \mathscr{P}_{m-1} \subset \ldots \subset \mathscr{P}_{i+1} \subset \mathscr{P}_i \subset \ldots \subset \mathscr{P}_1 \subset \mathscr{P}_0 = \mathscr{V}.$$

(Par définition $\mathbf{A}^0 = \mathbf{E}$, donc $\mathscr{P}_0 = \mathscr{V}$ même pour $\mathbf{A} = 0$.)

On a par construction $\mathbf{A}(\mathscr{P}_i) = \mathscr{P}_{i+1}$, $0 \leqslant i < m$, ce qui entraîne en particulier $\mathscr{P}_{m-1} \subset \operatorname{Ker} \mathbf{A}$. Par conséquent, on a pour toute base

$$\mathbf{e}_1^{(m-1)}, \ldots, \mathbf{e}_{p_{m-1}}^{(m-1)}, \quad p_{m-1} = \dim \mathscr{P}_{m-1}, \tag{1}$$

de l'espace $\mathscr{P}_{m-1}$ les relations

$$\mathbf{A}\mathbf{e}_1^{(m-1)} = \mathbf{0}, \ldots, \mathbf{A}\mathbf{e}_{p_{m-1}}^{(m-1)} = \mathbf{0}. \tag{2}$$

Il existe de plus dans $\mathscr{P}_{m-2}$ des vecteurs

$$\mathbf{e}_1^{(m-2)}, \ldots, \mathbf{e}_{p_{m-1}}^{(m-2)}$$

tels que

$$\mathbf{A}\mathbf{e}_1^{(m-2)} = \mathbf{e}_1^{(m-1)}, \ldots, \mathbf{A}\mathbf{e}_{p_{m-1}}^{(m-2)} = \mathbf{e}_{p_{m-1}}^{(m-1)}. \tag{3}$$

Il se trouve que *les vecteurs*

$$\begin{gathered} \mathbf{e}_1^{(m-1)}, \ldots, \mathbf{e}_{p_{m-1}}^{(m-1)}, \\ \mathbf{e}_1^{(m-2)}, \ldots, \mathbf{e}_{p_{m-1}}^{(m-2)} \end{gathered} \tag{4}$$

de $\mathscr{P}_{m-2}$ *sont linéairement indépendants.* En effet, si

$$k_1\mathbf{e}_1^{(m-1)} + \ldots + k_p\mathbf{e}_p^{(m-1)} + l_1\mathbf{e}_1^{(m-2)} + \ldots + l_p\mathbf{e}_p^{(m-2)} = \mathbf{0},$$

avec $p = p_{m-1}$, alors on a en faisant agir $\mathbf{A}$ et compte tenu de (2) et (3):

$$l_1\mathbf{e}_1^{(m-1)} + \ldots + l_p\mathbf{e}_p^{(m-1)} = \mathbf{0},$$

égalité qui a lieu seulement pour $l_1 = 0, \ldots, l_p = 0$ (parce que (1) est une base du sous-espace $\mathcal{P}_{m-1}$). Dans ce cas

$$k_1\mathbf{e}_1^{(m-1)} + \ldots + k_p\mathbf{e}_p^{(m-1)} = 0,$$

et les mêmes considérations entraînent $k_1 = 0, \ldots, k_p = 0$. □

Aussi on complète les vecteurs (4) en une base

$$\mathbf{e}_1^{(m-1)}, \ldots, \mathbf{e}_{p_{m-1}}^{(m-1)},$$

$$(5) \qquad \mathbf{e}_1^{(m-2)}, \ldots, \mathbf{e}_{p_{m-1}}^{(m-2)}, \ldots, \mathbf{e}_{p_{m-2}}^{(m-2)}, \quad p_{m-2} = \dim \mathcal{P}_{m-2} - \dim \mathcal{P}_{m-1},$$

de l'espace $\mathcal{P}_{m-2}$, et on voit sans peine que les vecteurs ajoutés

$$(6) \qquad \mathbf{e}_{p_{m-1}+1}^{(m-2)}, \ldots, \mathbf{e}_{p_{m-2}}^{(m-2)}$$

sont pris dans le noyau de l'opérateur $\mathbf{A}$, i.e. on les prend *de façon qu'on respecte les égalités*

$$(7) \qquad \mathbf{A}\mathbf{e}_{p_{m-1}+1}^{(m-2)} = 0, \ldots, \mathbf{A}\mathbf{e}_{p_{m-2}}^{(m-2)} = 0.$$

En effet, les vecteurs (1) forment une base de $\mathcal{P}_{m-1} = \mathbf{A}(\mathcal{P}_{m-2})$, si bien qu'on a pour tout choix de (6) et tout $i = 1, \ldots, p_{m-2} - p_{m-1}$ une égalité de la forme

$$\mathbf{A}\mathbf{e}_{p+i}^{(m-2)} = x_i^1\mathbf{e}_1^{(m-1)} + \ldots + x_i^p\mathbf{e}_p^{(m-1)}, \quad \text{où} \quad p = p_{m-1}.$$

Aussi on vérifie les conditions (7) en remplaçant $\mathbf{e}_{p+1}^{(m-2)}$ par les vecteurs

$$\mathbf{e}_{p+i}^{(m-2)} - x_i^1\mathbf{e}_1^{(m-2)} - \ldots - x_i^p\mathbf{e}_p^{(m-2)}. \quad \square$$

Puisque $\mathbf{A}(\mathcal{P}_{m-3}) = \mathcal{P}_{m-2}$, il existe dans $\mathcal{P}_{m-3}$ des vecteurs

$$(8) \qquad \mathbf{e}_1^{(m-3)}, \ldots, \mathbf{e}_{p_{m-2}}^{(m-3)}$$

tels que

$$(9) \qquad \mathbf{A}\mathbf{e}_1^{(m-3)} = \mathbf{e}_1^{(m-2)}, \ldots, \mathbf{A}\mathbf{e}_{p_{m-2}}^{(m-3)} = \mathbf{e}_{p_{m-2}}^{(m-2)}.$$

Comme pour la famille (4), on montre que *les vecteurs* (5) *et* (8) *constituent une famille libre*. En effet, si l'on fait agir $\mathbf{A}$ sur une combinaison linéaire quelconque de ces vecteurs, on aboutit par suite de (2), (3), (7), et (9) à une combinaison linéaire de (5). Les coefficients correspondants sont donc nuls, ce qui ramène la combinaison à celle de vecteurs (1) et (7) du noyau. Ces vecteurs étant linéairement indépendants, les coefficients restants sont également nuls. □

On complète cette famille libre de façon à avoir la base

$$\mathbf{e}_1^{(m-1)}, \ldots, \mathbf{e}_{p_{m-1}}^{(m-1)},$$

$$\mathbf{e}_1^{(m-2)}, \ldots, \mathbf{e}_{p_{m-1}}^{(m-2)}, \ldots, \mathbf{e}_{p_{m-2}}^{(m-2)},$$

$$\mathbf{e}_1^{(m-3)}, \ldots, \mathbf{e}_{p_{m-1}}^{(m-3)}, \ldots, \mathbf{e}_{p_{m-2}}^{(m-3)}, \ldots, \mathbf{e}_{p_{m-3}}^{(m-3)},$$

et on montre par des raisonnements analogues au cas (6) qu'on choisit les vecteurs à ajouter

$$\mathbf{e}^{(m-3)}_{p_{m-2}+1}, \ldots, \mathbf{e}^{(m-3)}_{p_{m-3}}$$

dans le noyau de A, i.e. on respecte les égalités

$$\mathbf{A}\mathbf{e}^{(m-3)}_{p_{m-2}+1} = 0, \ldots, \mathbf{A}\mathbf{e}^{(m-3)}_{p_{m-3}} = 0.$$

On obtient de proche en proche une base de $\mathcal{P}_0 = \mathcal{V}$ telle que ses vecteurs forment le tableau

$$\begin{array}{l}
\mathbf{e}^{(m-1)}_1, \ldots, \mathbf{e}^{(m-1)}_{p_{m-1}}, \\
\mathbf{e}^{(m-2)}_1, \ldots \mathbf{e}^{(m-2)}_{p_{m-1}}, \ldots, \mathbf{e}^{(m-2)}_{p_{m-2}}, \\
\mathbf{e}^{(m-3)}_1, \ldots, \mathbf{e}^{(m-3)}_{p_{m-1}}, \ldots, \mathbf{e}^{(m-3)}_{p_{m-2}}, \ldots, \mathbf{e}^{(m-3)}_{p_{m-3}}, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\mathbf{e}^{(0)}_1, \ldots, \mathbf{e}^{(0)}_{p_{m-1}}, \ldots, \mathbf{e}^{(0)}_{p_{m-2}}, \ldots, \mathbf{e}^{(0)}_{p_{m-3}}, \ldots, \mathbf{e}^{(0)}_{p_0}
\end{array}$$

dont chaque élément monte, sous l'action de A, d'une ligne tout en restant dans la même colonne (tandis que les vecteurs de la ligne supérieure s'annulent).

Cette propriété signifie par définition que les vecteurs de chaque colonne engendrent un sous-espace invariant, la restriction de A à chaque sous-espace étant un opérateur cyclique (dont le vecteur e est manifestement le dernier élément de la colonne; voir leçon précédente). L'espace $\mathcal{V}$ tout entier est somme directe de ces sous-espaces invariants, si bien que la proposition 6 de la leçon 15 se trouve complètement démontrée. □

* * *

Plaçons-nous de nouveau dans le cas d'opérateurs arbitraires.

On définit le sous-espace propre $\mathcal{P}_\lambda$ comme le plus grand sous-espace de $\mathcal{V}$ sur lequel $\mathbf{A} - \lambda\mathbf{E}$ est nul. Par analogie, on introduit la

Définition 1. Le plus grand sous-espace invariant $\mathcal{R}_\lambda$ de $\mathcal{V}$ sur lequel l'opérateur $\mathbf{A} - \lambda\mathbf{E}$ est nilpotent s'appelle le *sous-espace principal* de A associé à la valeur propre λ.

La définition demande quelques éclaircissements.

Soit $\mathbf{x} \neq 0$ un vecteur de $\mathcal{V}$. Il est dit *vecteur principal* associé à la valeur propre λ s'il existe un entier $m \geqslant 0$ tel que

$$(\mathbf{A} - \lambda\mathbf{E})^m\mathbf{x} = 0.$$

Il est clair que le vecteur $k\mathbf{x}$, $k \in \mathbb{K}$ quelconque, est un vecteur principal relatif à λ (ou le vecteur nul). L'affirmation analogue est également juste (on le constate facilement) pour la somme des vecteurs principaux associés à une même valeur propre car $(\mathbf{A} - \lambda\mathbf{E})^m \times$

$\times (x_1 + x_2) = 0$, avec $m = \max(m_1, m_2)$, si $(A - \lambda E)^{m_1} x_1 = 0$ et $(A - \lambda E)^{m_2} x_2 = 0$. Autrement dit, l'ensemble des vecteurs principaux relatifs à une valeur propre donnée λ plus le vecteur nul constituent un sous-espace de $\mathcal{T}$. Il s'agit justement du sous-espace $\mathcal{R}_\lambda$ de la définition 1 parce qu'il est évidemment invariant par $A - \lambda E$, donc par A.

Il est clair que chaque fois qu'un sous-espace $\mathcal{R}$ de $\mathcal{T}$ est invariant par A, il l'est par tout opérateur $A - \mu E$, $\mu \in K$. En particulier, *$\mathcal{R}_\lambda$ est invariant par tout opérateur de la forme* $A - \mu E$, $\mu \in K$, ce qui définit l'opérateur

$$(A - \mu E)|_{\mathcal{R}_\lambda}. \tag{10}$$

Proposition 1. *Si* $\mu \neq \lambda$, *l'opérateur* (10) *est inversible.*

Démonstration. Si

$$(A - \mu E)\, x = 0,$$

où $x \in \mathcal{R}_\lambda$, alors $Ax = \mu x$, donc $(A - \lambda E)\, x = (\mu - \lambda)\, x$. On en déduit que, ou bien le vecteur x est nul, ou bien le nombre $\mu - \lambda$ est valeur propre de l'opérateur nilpotent

$$(A - \lambda E)|_{\mathcal{R}_\lambda}.$$

Il est connu qu'un opérateur nilpotent arbitraire a toutes ses valeurs propres nulles. Par hypothèse $\mu \neq \lambda$, si bien que $x = 0$. Ainsi, le noyau de l'opérateur (10) est réduit au vecteur nul, d'où l'inversibilité de (10) (leçon 14, prop. 2). □

Avec ces résultats, on démontre pour les espaces principaux un analogue de la proposition 1 de la leçon 15.

Soient $\lambda_1, \ldots, \lambda_m$ des valeurs propres distinctes de A et

$$\mathcal{R}_1 = \mathcal{R}_{\lambda_1}, \ldots, \mathcal{R}_m = \mathcal{R}_{\lambda_m}$$

les sous-espaces principaux associés.

Proposition 2. *La somme*

$$\mathcal{R}_1 + \ldots + \mathcal{R}_m$$

des sous-espaces $\mathcal{R}_1, \ldots, \mathcal{R}_m$ *est directe, i.e. l'égalité*

$$x_1 + \ldots + x_m = 0,$$

où $x_1 \in \mathcal{R}_1, \ldots, x_m \in \mathcal{R}_m$, *a lieu si et seulement si*

$$x_1 = 0, \ldots, x_m = 0.$$

Démonstration (cf. leçon 15, prop. 1). L'affirmation est évidente pour $m = 1$. Supposons-la vraie pour $m - 1$ sous-espaces principaux. Comme $x_m \in \mathcal{R}_m$, on trouve un nombre s tel que

$$(A - \lambda_m E)^s\, x = 0.$$

Aussi

$$y_1 + \ldots + y_{m+1} = 0,$$

où

$$y_1 = (A - \lambda_m E)^s x_1, \ldots, y_{m-1} = (A - \lambda_m E)^s x_{m-1}.$$

Les espaces $\mathscr{R}_1, \ldots, \mathscr{R}_{m-1}$ étant invariants par $A - \lambda_m E$, on a $y_1 \in \mathscr{R}_1, \ldots, y_{m-1} \in \mathscr{R}_{m-1}$, et l'hypothèse de récurrence entraîne

$$y_1 = 0, \ldots, y_{m-1} = 0.$$

Selon la proposition 1, l'opérateur $A - \lambda_m E$ est inversible sur $\mathscr{R}_1, \ldots, \mathscr{R}_{m-1}$, d'où

$$x_1 = 0, \ldots, x_{m-1} = 0$$

et, partant, $x_m = 0$. □

L'avantage des sous-espaces principaux devant les sous-espaces propres est révélé par la

Proposition 3. *Etant donnée une valeur propre λ quelconque de l'opérateur A, la dimension du sous-espace principal $\mathscr{R}_\lambda$ est égale à la multiplicité algébrique de λ:*

$$\dim \mathscr{R}_\lambda = n_\lambda.$$

Démonstration. Soient $A_1 = A \mid_{\mathscr{R}_\lambda}$ et B l'opérateur de $\mathscr{V}/\mathscr{R}_\lambda$ dans $\mathscr{V}/\mathscr{R}_\lambda$ induit par A. Alors $f_A = f_{A_1} f_B$. Si $\dim \mathscr{R}_\lambda < n_\lambda$, le nombre λ est donc racine du polynôme f_B. Comme $\lambda \in K$, λ est encore une valeur propre de B.

Soit $x_0 + \mathscr{R}_\lambda$ un vecteur propre correspondant. On a

$$Ax_0 = \lambda x_0 + a_0,$$

où $a_0 \in \mathscr{R}_\lambda$, $a_0 = (A - \lambda E) x_0$. Il existe donc m tel que $(A - \lambda E)^m a_0 = 0$, auquel cas

$$(A - \lambda E)^{m+1} x_0 = 0,$$

donc $x_0 \in \mathscr{R}_\lambda$: contradiction qui démontre l'égalité $\dim \mathscr{R}_\lambda = n_\lambda$. □

La proposition 3 implique que si l'opérateur A a son spectre dans K et ne possède pas d'autres valeurs propres (racines caractéristiques) à part $\lambda_1, \ldots, \lambda_m$, alors

$$\dim (\mathscr{R}_{\lambda_1} \oplus \ldots \oplus \mathscr{R}_{\lambda_m}) = n_{\lambda_1} + \ldots + n_{\lambda_m} = n,$$

donc

$$\mathscr{R}_{\lambda_1} \oplus \ldots \oplus \mathscr{R}_{\lambda_m} = \mathscr{V}.$$

Ainsi, on a le

Théorème 1. *Etant donné un opérateur linéaire $A: \mathscr{V} \to \mathscr{V}$ quel-*

conque à spectre dans $\mathbb{K}$, *l'espace* $\mathcal{T}$ *est somme directe des sous-espaces principaux de* **A** :

$$\mathcal{T} = \mathcal{R}_{\lambda_1} \oplus \ldots \oplus \mathcal{R}_{\lambda_m}. \quad \square \tag{11}$$

* * *

Dire que le sous-espace invariant $\mathcal{R}$ de $\mathcal{T}$ est le sous-espace principal $\mathcal{R}_\lambda$, c'est dire que la restriction de **A** à $\mathcal{R}$ est la somme $\lambda\mathbf{E} + \mathbf{B}$ de l'opérateur scalaire $\lambda\mathbf{E}$ et d'un opérateur nilpotent **B**. Selon la proposition 6 de la leçon 15. l'espace $\mathcal{R}$ se décompose pour tout **B** en somme directe de sous-espaces invariants (par **B**, donc par **A**) sur chacun desquels **B** induit un opérateur cyclique. Si l'on effectue cette décomposition pour tout sous-espace principal, (11) on obtient $\mathcal{T}$ sous forme de somme directe de sous-espaces invariants tels que **A** induise sur chacun d'eux un opérateur

$$\lambda\mathbf{E} + \mathbf{C}, \tag{12}$$

avec $\lambda \in \mathbb{K}$ et **C** un opérateur cyclique.

Définition 2. Une matrice de la forme

$$\begin{pmatrix} \lambda & 1 & 0 & \ldots & \ldots & 0 \\ 0 & \lambda & 1 & 0 & \ldots & 0 \\ 0 & 0 & \lambda & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & \lambda & 1 \\ 0 & 0 & \ldots & \ldots & . & \lambda \end{pmatrix} \tag{13}$$

constitue un *bloc de Jordan*. Une matrice A somme directe de blocs de Jordan (à λ distincts en général) est appelée *forme normale de Jordan*.

Puisque tout opérateur cyclique est représenté dans une certaine base par la matrice (11) de la leçon 15, la matrice de l'opérateur (12) dans cette base est le bloc de Jordan (13). La réunion de toutes les bases des sous-espaces correspondants constitue donc une base de $\mathcal{T}$ dans laquelle la matrice de **A** s'écrit sous forme de Jordan. Cela démontre le

Théorème 2 (réduction à la forme de Jordan). *Etant donné un opérateur linéaire* $\mathbf{A} : \mathcal{T} \to \mathcal{T}$ *quelconque à spectre dans* $\mathbb{K}$, *il existe dans* $\mathcal{T}$ *une base par rapport à laquelle la matrice de* **A** *s'écrit sous forme normale de Jordan.* $\square$

Il se trouve que *la forme normale de Jordan de la matrice de l'opérateur est définie de façon unique* à l'ordre des blocs près, i.e. le nombre de blocs de Jordan, leur dimension et λ correspondants sont les mêmes pour toutes les bases dans lesquelles la matrice de l'opérateur a la forme normale. L'affirmation est évidente en ce qui con-

cerne λ (car ce sont valeurs propres de l'opérateur). Quant aux deux propriétés restantes, nous passons outre.

Remarque. L'unicité de la forme normale de Jordan d'une matrice découle de suite du fait que les blocs de Jordan d'ordre k associés à une valeur propre λ sont au nombre

$$r(A-\lambda E)^{k-1}-2r(A-\lambda E)^{k}+r(A-\lambda E)^{k+1},$$

$r(C)$ étant le rang de la matrice C.

La démonstration de la formule est un exercice utile que nous laissons au lecteur.

On souligne que la condition imposée au spectre de $\mathbf{A}$ par le dernier théorème est remplie pour $\mathbb{K} = \mathbb{C}$ par tout opérateur, si bien que *chaque opérateur linéaire sur* $\mathbb{C}$ *se réduit à la forme de Jordan.* □

* * *

Voyons un cas d'application des résultats obtenus.

Soit

$$f(x)=a_0x^m+a_1x^{m-1}+\ldots+a_m$$

un polynôme quelconque sur $\mathbb{K}$, auquel cas on définit pour tout $\mathbf{A}$ (resp. A) l'opérateur

$$f(\mathbf{A})=a_0\mathbf{A}^m+a_1\mathbf{A}^{m-1}+\ldots+a_m\mathbf{E}$$

(resp. la matrice $f(A) = a_0A^m + a_1A^{m-1} + \ldots + a_mE$) appelé *polynôme de l'opérateur* $\mathbf{A}$ (resp. de la matrice A).

Il est évident que tout sous-espace $\mathcal{P} \subset \mathcal{V}$ invariant par $\mathbf{A}$ l'est par tout opérateur $f(\mathbf{A})$, et

$$f(\mathbf{A})|_{\mathcal{P}} = f(\mathbf{A}|_{\mathcal{P}}). \tag{14}$$

En particulier, on définit pour tout $\mathbf{A}$ l'opérateur

$$f_{\mathbf{A}}(\mathbf{A}),$$

$f_{\mathbf{A}}(\lambda) = \det(A - \lambda E)$ étant le polynôme caractéristique de $\mathbf{A}$. On se propose de calculer $f_{\mathbf{A}}(\mathbf{A})$.

Soit

$$\mathbf{A} = \lambda_0\mathbf{E} + \mathbf{C}, \tag{15}$$

$\mathbf{C}$ étant un opérateur cyclique. Dans ce cas $f_{\mathbf{A}}(\lambda) = (\lambda_0 - \lambda)^n$ et $\mathbf{C}^n = 0$. Aussi

$$f_{\mathbf{A}}(\mathbf{A}) = (\lambda_0\mathbf{E} - \mathbf{A})^n = (-\mathbf{C})^n = 0.$$

Soient maintenant $\mathbf{A}$ un opérateur quelconque (à spectre dans $\mathbb{K}$) et

$$\mathcal{V} = \mathcal{P}_1 \oplus \ldots \oplus \mathcal{P}_N$$

la décomposition de $\mathscr{V}$ en somme directe de sous-espaces invariants tels que la restriction $\mathbf{A}_i = \mathbf{A}|_{\mathscr{P}_i}$ soit de la forme (15). Il découle des résultats obtenus

$$f_{\mathbf{A}_i}(\mathbf{A}_i) = 0. \tag{16}$$

Mais l'on sait que chaque polynôme $f_{\mathbf{A}_i}$ est un diviseur de $f_{\mathbf{A}}$ (on établit de même facilement que $f_{\mathbf{A}}$ est un produit de polynômes $f_{\mathbf{A}_i}, \ldots, f_{\mathbf{A}_N}$). Aussi (16) entraîne

$$f_{\mathbf{A}}(\mathbf{A}_i) = 0.$$

Donc (voir (14))

$$f_{\mathbf{A}}(\mathbf{A})|_{\mathscr{P}_i} = 0, \quad i = 1, \ldots, N.$$

Ainsi l'opérateur $f_{\mathbf{A}}(\mathbf{A})$ possède la propriété suivante : sa restriction au sous-espace $\mathscr{P}_i$ est nulle pour tout $i = 1, \ldots, N$. Par conséquent, $f_{\mathbf{A}}(\mathbf{A})$ est nul sur la somme des $\mathscr{P}_i$, i.e. sur $\mathscr{V}$ tout entier.

Cela prouve le

Théorème 3 (Hamilton-Cayley). *Tout opérateur annule son polynôme caractéristique* :

$$f_{\mathbf{A}}(\mathbf{A}) = 0. \ \square$$

Le théorème a été démontré pour les opérateurs à spectre dans $\mathbb{K}$ et, en particulier, pour tout opérateur sur $\mathbb{C}$. Dans la leçon suivante, on établira ce résultat dans le cas $\mathbb{R}$ pour les opérateurs à spectre quelconque.

LEÇON 17

Complexifié d'un opérateur linéaire. — Sous-espaces propres associés aux racines caractéristiques. — Opérateurs dont les complexifiés sont diagonalisables.

Les résultats de la leçon précédente ont été obtenus sous l'hypothèse de A à spectre dans le corps de base $\mathbb{K}$. La condition est nécessairement remplie pour $\mathbb{K} = \mathbb{C}$, tandis qu'elle en limite sensiblement le domaine d'application pour $\mathbb{K} = \mathbb{R}$. Nous allons examiner les opérateurs non assujettis à cette restriction. Le souci de simplicité nous fait nous borner à $\mathbb{K} = \mathbb{R}$, cas intéressant du point de vue géométrique, encore que la généralisation à $\mathbb{K}$ absolument quelconque soit possible au prix de modifications peu importantes.

* * *

On rappelle (voir I.19) qu'on associe à tout espace vectoriel $\mathscr{V}$ sur $\mathbb{R}$ un espace vectorie $\mathscr{V}^{\mathbb{C}}$ sur $\mathbb{C}$ qu'on appelle *complexifié* de $\mathscr{V}$. Chaque vecteur $\mathbf{z}$ de $\mathscr{V}^{\mathbb{C}}$ est représenté de façon unique par

$$\mathbf{z} = \mathbf{x} + i\mathbf{y},$$

où $\mathbf{x} \in \mathscr{V}$ et $\mathbf{y} \in \mathscr{V}$. Aussi la formule

$$\mathbf{A}^{\mathbb{C}}(\mathbf{x} + i\mathbf{y}) = \mathbf{A}\mathbf{x} + iA\mathbf{y}$$

définit bien pour tout opérateur linéaire $\mathbf{A} : \mathscr{V} \to \mathscr{V}$ un opérateur $\mathbf{A}^{\mathbb{C}} : \mathscr{V}^{\mathbb{C}} \to \mathscr{V}^{\mathbb{C}}$. On a pour tout nombre $a + ib \in \mathbb{C}$ et tout vecteur $\mathbf{z} = \mathbf{x} + i\mathbf{y} \in \mathscr{V}^{\mathbb{C}}$:

$$(a + ib)(\mathbf{x} + i\mathbf{y}) = (a\mathbf{x} - b\mathbf{y}) + i(a\mathbf{y} + b\mathbf{x}),$$

si bien que

$$\begin{aligned}\mathbf{A}^{\mathbb{C}}((a + ib)(\mathbf{x} + i\mathbf{y})) &= \mathbf{A}(a\mathbf{x} - b\mathbf{y}) + i\mathbf{A}(a\mathbf{y} + b\mathbf{x}) = \\ &= (a\mathbf{A}\mathbf{x} - b\mathbf{A}\mathbf{y}) + i(a\mathbf{A}\mathbf{y} + b\mathbf{A}\mathbf{x}) = \\ &= (a + ib)(\mathbf{A}\mathbf{x} + i\mathbf{A}\mathbf{y}) = (a + ib)\mathbf{A}^{\mathbb{C}}(\mathbf{x} + i\mathbf{y}),\end{aligned}$$

de sorte que $\mathbf{A}^{\mathbb{C}}(c\mathbf{z}) = c\mathbf{A}^{\mathbb{C}}\mathbf{z}$. L'égalité

$$\mathbf{A}^{\mathbb{C}}(\mathbf{z}_1 + \mathbf{z}_2) = \mathbf{A}^{\mathbb{C}}\mathbf{z}_1 + \mathbf{A}^{\mathbb{C}}\mathbf{z}_2,$$

avec $z_1, z_2 \in \mathcal{V}^{\mathbb{C}}$ quelconques, est encore plus simple à vérifier. Ainsi, *l'opérateur* $A^{\mathbb{C}}$ *est linéaire.* □

Définition 1. L'opérateur $A^{\mathbb{C}}$ s'appelle le *complexifié* de l'opérateur A.

Il est déjà apparu (voir I.19, prop. 1 ; avec les notations de cette proposition, $(\mathcal{V}^{\mathbb{C}})^{\mathbb{R}} = \mathcal{V}$) que toute base $e_1, \ldots, e_n$ de $\mathcal{V}$ en est une de $\mathcal{V}^{\mathbb{C}}$. Il en découle que la matrice de $A^{\mathbb{C}}$ dans cette base (« réelle ») coïncide avec celle de A. Dans ce sens, *la complexification d'un opérateur en conserve la matrice.* En particulier, A et $A^{\mathbb{C}}$ possèdent le même polynôme caractéristique :

$$f_A(\lambda) = f_{A^{\mathbb{C}}}(\lambda). \tag{1}$$

Il en résulte entre autres *la validité du théorème de Hamilton-Cayley* (leçon 16, th. 3) *pour tout opérateur* A *de* $\mathcal{V}$ *dans* $\mathcal{V}$. En effet, il joue, d'après la leçon 16, pour $A^{\mathbb{C}}$. L'opérateur $f_A(A)$ est évidemment la restriction de $f_A(A^{\mathbb{C}}) = f_{A^{\mathbb{C}}}(A^{\mathbb{C}})$ à $\mathcal{V} = \mathrm{Re}\,\mathcal{V}^{\mathbb{C}}$, si bien que $f_{A^{\mathbb{C}}}(A^{\mathbb{C}}) = 0$ implique $f_A(A) = 0$. □

Etant donnée l'égalité (1), A et $A^{\mathbb{C}}$ ont mêmes racines caractéristiques. L'opérateur $A^{\mathbb{C}}$ a toutes ses racines caractéristiques pour valeurs propres, et seules les racines caractéristiques réelles sont valeurs propres de A.

Si A est nilpotent, il en est de même de $A^{\mathbb{C}}$ (l'ordre de nilpotence étant le même pour A et $A^{\mathbb{C}}$) qui a donc toutes ses valeurs propres nulles.

Puisqu'il y a autant de valeurs propres de $A^{\mathbb{C}}$ que de racines du polynôme $f_{A^{\mathbb{C}}} = f_A$, on a $f_A(\lambda) = (-1)^n \lambda^n$ pour tout A nilpotent. Cela signifie en termes de matrices (avec le changement $\lambda = -\lambda$) que *toute matrice nilpotente A vérifie l'identité*

$$\det(A + \lambda E) = \lambda^n. \quad \square$$

Toute tentative de démontrer cette affirmation « purement matricielle » par un calcul direct du déterminant se heurte évidemment à de grandes difficultés.

Il en résulte par Hamilton-Cayley que *dans un espace de dimension n l'ordre de nilpotence d'un opérateur nilpotent arbitraire est égal au plus à n.* □

Ces affirmations élégantes montrent qu'en dépit de son apparence banale, la complexification constitue un outil excellent dans la démonstration des théorèmes.

Nous allons nous en servir pour étudier les racines caractéristiques.

* * *

Soit $\mathcal{Q}$ un sous-espace quelconque de $\mathcal{V}^{\mathbb{C}}$. Le sous-ensemble Re $\mathcal{Q}$ de tous les vecteurs de $\mathcal{V}$ de la forme Re $\mathbf{z}$, avec $\mathbf{z} \in \mathcal{Q}$, ou, ce qui revient au même (car Im $\mathbf{z}$ = Re $(-i\mathbf{z})$) de la forme Im $\mathbf{z}$, avec $\mathbf{z} \in \mathcal{Q}$, constitue évidemment un sous-espace de $\mathcal{V}$ (si $\mathbf{x}_1 = \mathrm{Re}\,\mathbf{z}_1$, $\mathbf{x}_2 = \mathrm{Re}\,\mathbf{z}_2$, alors $\mathbf{x}_1 + \mathbf{x}_2 = \mathrm{Re}\,(\mathbf{z}_1 + \mathbf{z}_2)$ et $k\,\mathrm{Re}\,\mathbf{z} = \mathrm{Re}\,k\mathbf{z}$ quel que soit $\mathbf{k} \in \mathbb{R}$).

De même, soit $\mathcal{P}$ un sous-espace quelconque de $\mathcal{V}$. L'ensemble $\mathcal{P}^{\mathbb{C}}$ formé de tous les vecteurs $\mathbf{x} + i\mathbf{y}$, $\mathbf{x}$, $\mathbf{y} \in \mathcal{P}$, est un sous-espace de $\mathcal{V}^{\mathbb{C}}$ (ce n'est rien autre que l'enveloppe linéaire de $\mathcal{P}$ dans $\mathcal{V}^{\mathbb{C}}$). Il est clair que

$$\mathrm{Re}\,\mathcal{P}^{\mathbb{C}} = \mathcal{P}$$

pour tout $\mathcal{P} \subset \mathcal{V}$.

On note que toute base $\mathbf{e}_1, \ldots, \mathbf{e}_p$ de $\mathcal{P}$ (sur $\mathbb{R}$) en est une de $\mathcal{P}^{\mathbb{C}}$ (sur $\mathbb{C}$).

Soient $\mathbf{A}: \mathcal{V} \to \mathcal{V}$ un opérateur linéaire quelconque agissant sur $\mathcal{V}$ et $\mathbf{A}^{\mathbb{C}}: \mathcal{V}^{\mathbb{C}} \to \mathcal{V}^{\mathbb{C}}$ son complexifié. Soit λ une racine caractéristique quelconque de $\mathbf{A}$. Elle est valeur propre de $\mathbf{A}^{\mathbb{C}}$, ce qui définit dans $\mathcal{V}^{\mathbb{C}}$ le sous-espace propre associé $\mathcal{Q}_\lambda$.

Supposons λ réelle, auquel cas elle est valeur propre de $\mathbf{A}$, et on définit dans $\mathcal{V}$ le sous-espace propre correspondant $\mathcal{P}_\lambda$.

Etant donné un système de n équations linéaires homogènes à n inconnues dont les coefficients sont réels et forment une matrice de rang r, les solutions du système constituen tdans $\mathbb{R}^n$ un sous-espace $\mathcal{P}$ de dimension $n - r$, et chaque solution est combinaison linéaire de certaines $n - r$ solutions linéairement indépendantes formant une base de $\mathcal{P}$. On rappelle que cette base est par définition le *système fondamental de solutions.*

On considère le même système d'équations comme système à coefficients complexes, auquel cas on cherche les solutions dans $\mathbb{C}^n = (\mathbb{R}^n)^{\mathbb{C}}$. Chaque système fondamental de solutions reste alors un système fondamental, mais on prend les combinaisons linéaires de solutions précédées de coefficients complexes arbitraires et non de coefficients réels. Avec les notations introduites plus haut, cela veut dire que les solutions du système donné forment dans $\mathbb{C}^n$ un sous-espace $\mathcal{P}^{\mathbb{C}}$.

Ces considérations générales s'appliquent en particulier au sous-espace $\mathcal{P}_\lambda$ de vecteurs dont les coordonnées dans une base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ quelconque de $\mathcal{V}$ vérifient un système d'équations linéaires homogènes à matrice réelle des coefficients $A - \lambda E$. On sait déjà que les vecteurs $\mathbf{e}_1, \ldots, \mathbf{e}_n$ forment aussi une base de $\mathcal{V}^{\mathbb{C}}$ et que les coordonnées des vecteurs de $\mathcal{Q}_\lambda$ dans cette base se définissent par le même système d'équations. Ainsi, *toute racine caractéristique réelle λ de*

l'opérateur **A** *satisfait à*

$$\mathcal{G}_\lambda = \mathcal{P}_\lambda^{\mathbb{C}},$$

donc à

$$\mathcal{P}_\lambda = \operatorname{Re} \mathcal{G}_\lambda. \tag{2}$$

Voyons ce qu'il en est de λ non réelle. Dans ce cas, on définit $\mathcal{P}_\lambda \subset \mathcal{T}$ par la formule (2).

Ainsi, les sous-espaces $\mathcal{P}_\lambda$ sont définis pour toute racine caractéristique λ de **A**, la notation ayant pour λ réelle le sens ci-dessus.

Nous dirons du sous-espace $\mathcal{P}_\lambda = \operatorname{Re} \mathcal{G}_\lambda$ que c'est *le sous-espace propre de* **A** *associé à la racine caractéristique* λ. On n'oubliera pas que ses éléments sont vecteurs propres de **A** pour λ réelle seule.

Il est clair que chaque $\mathcal{P}_\lambda$ est invariant par **A**.

Si λ est réelle, on a $\mathcal{P}_\lambda^{\mathbb{C}} = \mathcal{G}_\lambda$. Que devient $\mathcal{P}_\lambda^{\mathbb{C}}$ pour λ non réelle ?

On note au préalable la propriété suivante du polynôme caractéristique f_A à coefficients réels : chaque fois qu'il admet λ pour racine, il a également comme racine le complexe conjugué $\bar{\lambda}$. Les coordonnées des vecteurs du sous-espace propre correspondant $\mathcal{G}_{\bar{\lambda}}$ de l'opérateur $\mathbf{A}^{\mathbb{C}}$ sont solutions d'un système d'équations linéaires de matrice complexe conjuguée $A - \bar{\lambda}E$, si bien qu'on les obtient à partir des coordonnées des éléments de $\mathcal{G}_\lambda$ par passage aux nombres complexes conjugués. Cela signifie, en faisant abstraction des coordonnées, que si $\mathbf{z} = \mathbf{x} + i\mathbf{y} \in \mathcal{G}_\lambda$, alors $\bar{\mathbf{z}} = \mathbf{x} - i\mathbf{y} \in \mathcal{G}_{\bar{\lambda}}$. Ce fait s'écrit de façon conventionnelle mais intuitive :

$$\mathcal{G}_{\bar{\lambda}} = \bar{\mathcal{G}}_\lambda.$$

D'où $\operatorname{Re} \mathcal{G}_{\bar{\lambda}} = \operatorname{Re} \mathcal{G}_\lambda$, i.e.

$$\mathcal{P}_\lambda = \mathcal{P}_{\bar{\lambda}}.$$

D'autre part, la somme des sous-espaces $\mathcal{G}_\lambda$ et $\mathcal{G}_{\bar{\lambda}}$ est la somme directe $\mathcal{G}_\lambda \oplus \mathcal{G}_{\bar{\lambda}}$ du moment que $\lambda \neq \bar{\lambda}$ (leçon 15, prop. 1). Si $\mathbf{e}_1, \ldots, \mathbf{e}_q$ est une base de $\mathcal{G}_\lambda$, où $q = \dim \mathcal{G}_\lambda$, les vecteurs $\mathbf{e}_1, \ldots, \mathbf{e}_q, \bar{\mathbf{e}}_1, \ldots, \bar{\mathbf{e}}_q$ constituent une base de l'espace $\mathcal{G}_\lambda \oplus \mathcal{G}_{\bar{\lambda}}$. Une autre base de cet espace est alors formée par les vecteurs

$$\operatorname{Re} \mathbf{e}_1 = \frac{\mathbf{e}_1 + \bar{\mathbf{e}}_1}{2}, \qquad \operatorname{Im} \mathbf{e}_1 = \frac{\mathbf{e}_1 - \bar{\mathbf{e}}_1}{2i},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\operatorname{Re} \mathbf{e}_q = \frac{\mathbf{e}_q + \bar{\mathbf{e}}_q}{2}, \qquad \operatorname{Im} \mathbf{e}_q = \frac{\mathbf{e}_q - \bar{\mathbf{e}}_q}{2i}.$$

Les vecteurs $\operatorname{Re} \mathbf{e}_1, \operatorname{Im} \mathbf{e}_1, \ldots, \operatorname{Re} \mathbf{e}_q, \operatorname{Im} \mathbf{e}_q$ sont réels par construction, i.e. ils sont dans $\mathcal{T}$. Si $\mathcal{P} \subset \mathcal{T}$ est le sous-espace de

$\mathcal{T}$ de dimension $2q$ engendré par ces vecteurs, on a par définition

$$\mathcal{T}^{\mathbb{C}} = \mathcal{G}_\lambda \otimes \mathcal{G}_{\bar\lambda}.$$

On vérifie cependant sans peine que la correspondance $\mathcal{G} \longmapsto \operatorname{Re} \mathcal{G}$ transforme la somme de sous-espaces en une somme, i.e.

$$\operatorname{Re}(\mathcal{G}_1 + \mathcal{G}_2) = \operatorname{Re} \mathcal{G}_1 + \operatorname{Re} \mathcal{G}_2$$

quels que soient $\mathcal{G}_1, \mathcal{G}_2 \subset \mathcal{T}^{\mathbb{C}}$. Aussi

$$\mathcal{P} = \operatorname{Re} \mathcal{P}^{\mathbb{C}} = \operatorname{Re}(\mathcal{G}_\lambda \oplus \mathcal{G}_{\bar\lambda}) = \operatorname{Re} \mathcal{G}_\lambda + \operatorname{Re} \mathcal{G}_{\bar\lambda} = \mathcal{P}_\lambda + \mathcal{P}_{\bar\lambda} = \mathcal{P}_\lambda.$$

Cela démontre que *toute racine caractéristique non réelle λ de l'opérateur* A *satisfait à l'égalité*

$$\mathcal{P}_\lambda^{\mathbb{C}} = \mathcal{G}_\lambda \oplus \mathcal{G}_{\bar\lambda}. \quad \square$$

Avec les sous-espaces $\mathcal{G}_1 = \mathcal{G}_\lambda$ et $\mathcal{G}_2 = \mathcal{G}_{\bar\lambda}$, la correspondance $\mathcal{G} \longmapsto \operatorname{Re} \mathcal{G}$ ne transforme pas nécessairement la somme directe en une somme directe. Mais il est immédiat d'établir que *si* $\mathcal{G}_1 = \mathcal{P}_1^{\mathbb{C}}$ *et* $\mathcal{G}_2 = \mathcal{P}_2^{\mathbb{C}}$, *alors*

$$\operatorname{Re}(\mathcal{G}_1 \oplus \mathcal{G}_2) = \operatorname{Re} \mathcal{G}_1 \oplus \operatorname{Re} \mathcal{G}_2. \tag{3}$$

En effet, il est clair que

$$(\mathcal{P}_1 \oplus \mathcal{P}_2)^{\mathbb{C}} = \mathcal{P}_1^{\mathbb{C}} \oplus \mathcal{P}_2^{\mathbb{C}}.$$

Aussi on obtient (3) par application de Re. $\square$

On se propose de démontrer pour $\mathcal{P}_\lambda$ un analogue de la proposition 1 de la leçon 15.

Proposition 1. *Soient $\lambda_1, \ldots, \lambda_m$ des racines caractéristiques (réelles ou non) de l'opérateur* A *telles que*

$$\lambda_i \neq \lambda_j \quad \text{et} \quad \lambda_i \neq \bar\lambda_j. \qquad i, j = 1, \ldots, m,$$

pour $i \neq j$. La somme $\mathcal{P}$ des sous-espaces $\mathcal{P}_{\lambda_1}, \ldots, \mathcal{P}_{\lambda_m}$ en est somme directe:

$$\mathcal{P} = \mathcal{P}_{\lambda_1} \oplus \ldots \oplus \mathcal{P}_{\lambda_m}. \tag{4}$$

Le cas des sous-espaces $\mathcal{P}_\lambda$ et $\mathcal{P}_{\bar\lambda}$ témoigne du caractère essentiel de la condition $\lambda_i \neq \bar\lambda_j$.

Démonstration. Soient $\lambda_1, \ldots, \lambda_r$ les racines caractéristiques réelles et $\lambda_{r+1}, \ldots, \lambda_m$ les racines non réelles. Alors $2m - r$ valeurs propres

$$\lambda_1, \ldots, \lambda_r, \lambda_{r+1}, \bar\lambda_{r+1}, \ldots, \lambda_m, \bar\lambda_m$$

de l'opérateur $A^{\mathbb{C}}$ sont toutes distinctes, si bien que la somme $\mathcal{Q}$ des sous-espaces propres associés est somme directe :

$$\mathcal{Q} = \mathcal{Q}_{\lambda_1} \oplus \cdots \oplus \mathcal{Q}_{\lambda_r} \oplus (\mathcal{Q}_{\lambda_{r+1}} \oplus \mathcal{Q}_{\bar{\lambda}_{r+1}}) \oplus \cdots \oplus (\mathcal{Q}_{\lambda_m} \oplus \mathcal{Q}_{\bar{\lambda}_m}).$$

On applique Re et on tient compte de

$$\mathcal{Q}_{\lambda_1} = \mathcal{P}^{\mathbb{C}}_{\lambda_1}, \ldots, \mathcal{Q}_{\lambda_r} = \mathcal{P}^{\mathbb{C}}_{\lambda_r},$$

$$\mathcal{Q}_{\lambda_{r+1}} \oplus \mathcal{Q}_{\bar{\lambda}_{r+1}} = \mathcal{P}^{\mathbb{C}}_{\lambda_{r+1}}, \ldots, \mathcal{Q}_{\lambda_m} \oplus \mathcal{Q}_{\bar{\lambda}_m} = \mathcal{P}^{\mathbb{C}}_{\lambda_m},$$

il vient de suite (4). □

* * *

Si $A^{\mathbb{C}}$ est diagonalisable, la proposition 1 entraîne que $\mathcal{T}$ est somme directe de $\mathcal{P}_\lambda$ invariants (par A), λ parcourant toutes les racines réelles du polynôme f_A et toutes ses racines non réelles deux à deux non conjuguées.

On connaît pour λ réel la restriction $A_\lambda = A|_{\mathcal{P}_\lambda}$ de l'opérateur A au sous-espace $\mathcal{P}_\lambda$. C'est λE, opérateur scalaire de matrice λE qui est diagonale (scalaire) dans toute base.

Soit λ non réel. Voyons ce qu'il en devient de l'opérateur $A_\lambda = A|_{\mathcal{P}_\lambda}$. Soit

$$\lambda = \alpha + i\beta, \text{ où } \alpha, \beta \in \mathbb{R} \text{ et } \beta \neq 0.$$

On a déjà montré que quelle que soit la base $e_1, \ldots, e_q$ du sous-espace $\mathcal{Q}_\lambda$, les vecteurs

$$\text{(5)} \qquad \operatorname{Re} e_1, \operatorname{Im} e_1, \ldots, \operatorname{Re} e_q, \operatorname{Im} e_q$$

constituent une base de $\mathcal{P}_\lambda$. On pose $e = e_1$, $x = \operatorname{Re} e$, $y = \operatorname{Im} e$ (pour simplifier les notations), et on considère un sous-espace $\mathcal{P} \subset \mathcal{P}_\lambda$ de dimension 2 qui admet x, y pour base.

Comme $e \in \mathcal{Q}_\lambda$, on a $A^{\mathbb{C}}e = \lambda e$, i.e.

$$A^{\mathbb{C}}(x + iy) = (\alpha + i\beta)(x + iy).$$

Cela signifie par définition que

$$Ax + iAy = (\alpha x - \beta y) + i(\beta y + \alpha y),$$

i.e.

$$Ax = \alpha x - \beta y,$$
$$Ay = \beta x + \alpha y.$$

Ainsi, l'*espace* $\mathcal{P}$ *est invariant par l'opérateur* A, *et la restriction de* A *à* $\mathcal{P}$ *a pour matrice dans la base* x, y

$$\text{(6)} \qquad \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}. \quad \square$$

L'espace $\mathcal{P}_\lambda$ étant somme directe de q sous-espaces $\mathcal{P}$, la matrice de A_λ dans la base (5) est bloc-diagonale

$$\begin{pmatrix} \begin{matrix} \alpha_1 & \beta_1 \\ -\beta_1 & \alpha_1 \end{matrix} & & & 0 \\ & \begin{matrix} \alpha_2 & \beta_2 \\ -\beta_2 & \alpha_2 \end{matrix} & & \\ & & \ddots & \\ 0 & & & \begin{matrix} \alpha_q & \beta_q \\ -\beta_q & \alpha_q \end{matrix} \end{pmatrix}$$

dont les éléments diagonaux sont $q = \dim \mathcal{Q}_\lambda$ matrices (6).

Le bilan des résultats obtenus s'énonce sous forme de

Théorème 1. *Etant donné un opérateur linéaire* $A : \mathcal{V} \to \mathcal{V}$ *sur* $\mathbb{R}$, *on suppose que l'opérateur* $A^{\mathbb{C}} : \mathcal{V}^{\mathbb{C}} \to \mathcal{V}^{\mathbb{C}}$ *est diagonalisable (il l'est en particulier pour* A *à spectre simple). Soient* $\lambda_1, \ldots, \lambda_r$ *les racines réelles de* A *et* $\lambda_{r+1} = \alpha_1 + i\beta_1, \ldots, \lambda_m = \alpha_{m-r} + i\beta_{m-r}$ *ses racines non réelles deux à deux complexes non conjuguées, chacune d'elles étant comptée son ordre de multiplicité (si bien que* $m = 2n - r$).

Il existe dans l'espace $\mathcal{V}$ *une base dans laquelle la matrice de l'opérateur* A *est somme directe d'un certain nombre de matrices* $\lambda_1, \ldots, \lambda_r$ *d'ordre* 1 *et de matrices*

$$\begin{pmatrix} \alpha_1 & \beta_1 \\ -\beta_1 & \alpha_1 \end{pmatrix}, \ldots, \begin{pmatrix} \alpha_{m-r} & \beta_{m-r} \\ -\beta_{m-r} & \alpha_{m-r} \end{pmatrix}$$

d'ordre 2, *i.e. elle est de la forme*

$$(7) \qquad \begin{pmatrix} \lambda_1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & \lambda_r & & & & & & \\ & & & \ddots & & & & & \\ & & & & \begin{matrix} \alpha_1 & \beta_1 \\ -\beta_1 & \alpha_1 \end{matrix} & & & & \\ & & & & & \ddots & & & \\ 0 & & & & & & \begin{matrix} \alpha_{m-r} & \beta_{m-r} \\ -\beta_{m-r} & \alpha_{m-r} \end{matrix} \end{pmatrix}. \quad \square$$

On conçoit que l'opérateur $A^{\mathbb{C}}$ non diagonalisable donne lieu à un théorème analogue avec la matrice (7) plus complexe. Nous nous abstiendrons de le formuler et de le démontrer car il ne nous sera d'aucune utilité.

LEÇON 18

Espaces euclidiens et hermitiens. — Supplémentaires orthogonaux. — Identification de vecteurs et de covecteurs. — Annulateurs et supplémentaires orthogonaux. — Fonctionnelles bilinéaires et opérateurs linéaires. — Une règle imposée à l'identification de tenseurs de différents types. — Tenseur métrique. — Elévation et abaissement d'indices.

Un espace vectoriel $\mathcal{V}$ sur le corps $\mathbb{R}$ des nombres réels est *euclidien* s'il est muni d'une fonctionnelle bilinéaire symétrique définie positive. On dit que la fonctionnelle définit sur l'espace un *produit scalaire*. La valeur de la fonctionnelle sur deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ s'appelle *le produit scalaire de* $\mathbf{x}$ *et* $\mathbf{y}$ et se note $\mathbf{xy}$ ou encore $(\mathbf{x}, \mathbf{y})$ (I.13, déf. 2 et 3).

La notion de nombre positif étant absente de $\mathbb{C}$, on ne saurait transposer directement ces concepts à ce corps. On prend donc des chemins détournés.

Définition 1. Une fonctionnelle $\mathbf{x}, \mathbf{y} \to B(\mathbf{x}, \mathbf{y})$ définie sur un espace vectoriel complexe $\mathcal{V}$ est dite *sesquilinéaire* si elle est linéaire par rapport à la première variable, i.e. si

$$B(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = B(\mathbf{x}_1, \mathbf{y}) + B(\mathbf{x}_2, \mathbf{y})$$

et

$$B(c\mathbf{x}, \mathbf{y}) = cB(\mathbf{x}, \mathbf{y})$$

quels que soient $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}, \mathbf{y} \in \mathcal{V}$ et $c \in \mathbb{C}$, et *semi-linéaire* par rapport à la seconde, i.e. si

$$B(\mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2) = B(\mathbf{x}, \mathbf{y}_1) + B(\mathbf{x}, \mathbf{y}_2)$$

et

$$B(\mathbf{x}, c\mathbf{y}) = \bar{c}B(\mathbf{x}, \mathbf{y})$$

quels que soient $\mathbf{x}, \mathbf{y}_1, \mathbf{y}_2, \mathbf{y} \in \mathcal{V}$ et $c \in \mathbb{C}$.

On dit d'une fonctionnelle sesquilinéaire B qu'elle est *hermitienne* si

$$B(\mathbf{y}, \mathbf{x}) = \overline{B(\mathbf{x}, \mathbf{y})}$$

pour tout $\mathbf{x}$ et tout $\mathbf{y}$ de $\mathcal{V}$.

S'agissant d'une fonctionelle hermitienne, le nombre $B(\mathbf{x}, \mathbf{x})$ est réel pour tout vecteur $\mathbf{x} \in \mathcal{V}$, si bien que la condition de son signe a un sens.

Une fonctionelle hermitienne B est par définition *définie positive* si

$$B(\mathbf{x}, \mathbf{x}) > 0$$

pour tout vecteur $\mathbf{x} \neq 0$ de l'espace $\mathcal{V}$.

Un espace vectoriel $\mathcal{V}$ sur $\mathbb{C}$ s'appelle *espace hermitien* (ou *unitaire*) s'il est muni d'une fonctionnelle sesquilinéaire hermitienne définie positive. On dit que la fonctionnelle définit sur l'espace un *produit scalaire*. La valeur de la fonctionnelle sur deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ est *le produit scalaire de* $\mathbf{x}$ et $\mathbf{y}$ et se note $\mathbf{xy}$ ou encore $(\mathbf{x}, \mathbf{y})$

L'exemple d'espace hermitien est l'espace $\mathbb{C}^n$ à produit scalaire défini par la formule

$$(\mathbf{x}, \mathbf{y}) = x_1\bar{y}_1 + \ldots + x_n\bar{y}_n.$$

Les éléments de base de la théorie des espaces hermitiens sont complètement analogues à ce que nous connaissons déjà de la théorie des espaces euclidiens (voir I.13; I.14). Dans un espace hermitien, on définit par exemple la *longueur* $|\mathbf{x}| = \sqrt{(\mathbf{x}, \mathbf{x})}$ de tout vecteur $\mathbf{x}$; l'inégalité de Cauchy-Bouniakovski est juste, ainsi que l'inégalité de Bessel (I.14, prop. 2; x_i^2 fait naturellement place à $|x_i|^2$); les notions de *vecteurs orthogonaux* ($(\mathbf{x}, \mathbf{y}) = 0$) et de familles orthonormées (en particulier, la notion de *bases orthonormées*) ont un sens; on a un analogue de la proposition 3 de I.14 relative aux propriétés des bases orthonormées (mais l'égalité de Parseval s'écrit p. ex. $(\mathbf{x}, \mathbf{y}) = x_1\bar{y}_1 + \ldots + x_n\bar{y}_n$); on applique le procédé d'orthonormalisation de Gram-Schmidt, et ainsi de suite. On conçoit que deux théorèmes homologues ont une interprétation géométrique différente selon qu'il s'agit d'un espace euclidien ou d'un espace hermitien. S'agissant d'espaces euclidiens, l'existence d'une base orthonormée signifie par exemple que tout espace euclidien de dimension n est isomorphe à $\mathbb{R}^n$ muni du produit scalaire $(\mathbf{x}, \mathbf{y}) = x_1y_1 + \ldots + x_ny_n$, tandis qu'elle équivaut pour les espaces hermitiens à l'isomorphisme entre tout espace hermitien n-dimensionnel et l'espace $\mathbb{C}^n$ à produit scalaire défini par l'égalité $(\mathbf{x}, \mathbf{y}) = x_1\bar{y}_1 + \ldots + x_n\bar{y}_n$.

Dans la suite, nous essayerons d'énoncer les théorèmes pour les deux catégories d'espaces à la fois, et nous désignerons de préférence le produit scalaire par $(\mathbf{x}, \mathbf{y})$ (contrairement au livre I).

* * *

Soit S une partie quelconque d'un espace vectoriel $\mathcal{V}$ euclidien ou unitaire.

Définition 2. On appelle *supplémentaire orthogonal* de S, et on note $S^\perp$, l'ensemble des vecteurs de $\mathcal{V}$ orthogonaux à chaque vecteur de S:

$$S^\perp = \{\mathbf{y} \in \mathcal{V};\ (\mathbf{x}, \mathbf{y}) = 0,\ \mathbf{x} \in S\}.$$

Les propriétés des supplémentaires orthogonaux sont analogues à celles des annulateurs (voir leçon 4). Il est clair par exemple que *le supplémentaire orthogonal de tout ensemble est un sous-espace et* $S^{\perp} \supset T^{\perp}$ *si* $S \subset T$. On a l'analogue suivant de la proposition 5 de la leçon 4.

Proposition 1. *Le supplémentaire orthogonal d'un ensemble S quelconque est identique au supplémentaire orthogonal de son enveloppe linéaire*:

$$S^{\perp} = [S]^{\perp}.$$

Démonstration. Comme $S \subset [S]$, on a $S^{\perp} \supset [S]^{\perp}$. Inversement, si $\mathbf{y} \in S^{\perp}$ et $\mathbf{x} = k_1\mathbf{x}_1 + \ldots + k_m\mathbf{x}_m$, où $\mathbf{x}_1, \ldots, \mathbf{x}_m \in S$, alors

$$(\mathbf{x}, \mathbf{y}) = k_1 (\mathbf{x}_1, \mathbf{y}) + \ldots + k_m (\mathbf{x}_m, \mathbf{y}) = 0,$$

donc $\mathbf{y} \in [S]^{\perp}$. □

On se limite donc sans restreindre la généralité aux supplémentaires orthogonaux des sous-espaces.

Par analogie avec la proposition 6 de la leçon 4, on établit un résultat portant sur la dimension des orthogonaux. Une affirmation plus forte a également lieu car le supplémentaire orthogonal de tout sous-espace $\mathcal{P} \subset \mathcal{V}$ est un sous-espace de $\mathcal{V}$.

Proposition 2. *Etant donné un sous-espace* $\mathcal{P} \subset \mathcal{V}$ *quelconque, l'espace* $\mathcal{V}$ *est somme directe de* $\mathcal{P}$ *et de son orthogonal*:

$$\mathcal{V} = \mathcal{P} \oplus \mathcal{P}^{\perp}.$$

Démonstration. Soient $\mathbf{e}_1, \ldots, \mathbf{e}_p$ une base orthonormée de $\mathcal{P}$ et $\mathbf{x} \in \mathcal{V}$. On désigne par $x_1, \ldots, x_p$ les coefficients de Fourier de $\mathbf{x}$ par rapport à la famille orthonormée $\mathbf{e}_1, \ldots, \mathbf{e}_p$, et on construit le vecteur $\mathbf{x}' = x_1\mathbf{e}_1 + \ldots + x_p\mathbf{e}_p$. D'après I.14, prop. 2, le vecteur $\mathbf{x} - \mathbf{x}'$ est orthogonal à tous les vecteurs $\mathbf{e}_1, \ldots, \mathbf{e}_p$ et appartient donc (prop. 1) au sous-espace $\mathcal{P}^{\perp}$. Ainsi,

$$\mathbf{x} = \mathbf{x}' + (\mathbf{x} - \mathbf{x}'),$$

avec $\mathbf{x}' \in \mathcal{P}$ et $\mathbf{x} - \mathbf{x}' \in \mathcal{P}^{\perp}$, ce qui démontre l'égalité $\mathcal{V} = \mathcal{P} + \mathcal{P}^{\perp}$.

Il reste à prouver que $\mathcal{P} \cap \mathcal{P}^{\perp} = 0$. C'est d'ailleurs évident parce qu'avec $\mathbf{x} \in \mathcal{P} \cap \mathcal{P}^{\perp}$, donc $(\mathbf{x}, \mathbf{y}) = 0$ pour tout $\mathbf{y} \in \mathcal{P}$, on a en particulier $(\mathbf{x}, \mathbf{x}) = 0$ et, partant, $\mathbf{x} = 0$. □

L'analogue de la proposition 1 de la leçon 5 vient tout naturellement, à savoir on a la

Proposition 3. *On a pour tout sous-espace* $\mathcal{P} \subset \mathcal{P}$:

$$\mathcal{P}^{\perp\perp} = \mathcal{P}.$$

Démonstration. Si $\mathbf{x} \in \mathcal{P}$, alors $(\mathbf{x}, \mathbf{y}) = 0$ pour tout $\mathbf{y} \in \mathcal{P}^{\perp}$, donc $\mathbf{x} \in \mathcal{P}^{\perp\perp}$. Par conséquent, $\mathcal{P} \subset \mathcal{P}^{\perp\perp}$. Inversement, soit $\mathbf{x} \in \mathcal{P}^{\perp\perp}$. On utilise la proposition 2 et on pose $\mathbf{x} = \mathbf{x}' + \mathbf{x}''$, avec $\mathbf{x}' \in \mathcal{P}$, $\mathbf{x}'' \in \mathcal{P}^{\perp}$, si bien que $(\mathbf{x}', \mathbf{x}'') = 0$. Puisque $\mathbf{x} \in \mathcal{P}^{\perp\perp}$, on a $(\mathbf{x}, \mathbf{x}'') =$

$= 0$, donc

$$(x'', x'') = (x - x', x'') = (x, x'') - (x', x'') = 0.$$

D'où $x'' = 0$, donc $x = x' \subset \mathcal{P}$. Ainsi, $\mathcal{P} = \mathcal{P}^{\perp\perp}$. □

On souligne que l'espace vectoriel peut être aussi bien euclidien qu'hermitien.

* * *

Ce parallélisme entre espaces euclidiens et espaces hermitiens ne concerne pas leurs duals. L'espace dual $\mathcal{V}'$ de $\mathcal{V}$ euclidien et son homologue $\mathcal{V}''$ pour $\mathcal{V}$ hermitien seront donc étudiés séparément.

Soit $\mathcal{V}$ un espace vectoriel euclidien.

Proposition 4. *Quel que soit $\mathcal{V}$ euclidien, on a l'isomorphisme canonique*

$$\mathcal{V}'' \approx \mathcal{V}.$$

Démonstration. Il suffit de démontrer que

$$\mathcal{V} | \mathcal{V}'$$

(leçon 4, prop. 4), i.e. qu'il existe un couplage naturel de $\mathcal{V}$ et $\mathcal{V}'$. Or, ce couplage existe effectivement. C'est un produit scalaire (cette fonctionnelle étant positive est évidemment non dégénérée). □

L'isomorphisme $\mathcal{V}' \to \mathcal{V}''$ se définit explicitement comme application qui associe à tout vecteur $y \in \mathcal{V}'$ la fonctionnelle linéaire

$$\xi_y : x \mapsto (x, y).$$

La correspondance $y \mapsto \xi_y$ est visiblement un homomorphisme. Les espaces $\mathcal{V}'$ et $\mathcal{V}''$ sont de même dimension, si bien qu'on établit la propriété, pour cet homomorphisme, d'être un isomorphisme sous la condition suffisante de nullité de son noyau ($\xi_y \neq 0$ si $y \neq 0$). La chose est évidente parce que, disons, $\xi_y(y) = (y, y) \neq 0$.

Nous avons en fait répété la démonstration de la proposition 4 de la leçon 4 pour le couplage $x, y \mapsto (x, y)$.

A première vue, la démonstration semble valide pour les espaces hermitiens. Une étude plus attentive permet cependant de constater que *l'application* $y \mapsto \xi_y$ *n'est pas un homomorphisme pour $\mathcal{V}'$ hermitien* car si elle change la somme en une somme:

$$\xi_{y_1+y_2} = \xi_{y_1} + \xi_{y_2}, \quad y_1, y_2 \in \mathcal{V},$$

tout produit par un nombre devient dans cette transformation un produit par le nombre complexe conjugué, i.e. on a pour tout vecteur $y \in \mathcal{V}'$ et tout nombre $c \in \mathbb{C}$:

$$\xi_{cy} = \bar{c}\xi_y.$$

Les applications d'espaces vectoriels sur $\mathbb{C}$ ayant ces propriétés sont dites *semi-linéaires*. On voit facilement qu'à l'instar des applications linéaires, les applications semi-linéaires d'espaces vectoriels de même dimension à noyau nul sont bijectives (ce sont des *isomorphismes semi-linéaires*). Aussi on répète les raisonnements de la proposition 4 pour montrer que *le dual $\mathcal{T}''$ de $\mathcal{T}'$ hermitien lui est semi-linéairement isomorphe.*

Ce résultat ne nous satisfait évidemment pas. En effet, l'intérêt capital de la proposition 4 est qu'elle permet d'*identifier* chaque espace euclidien $\mathcal{T}'$ et son dual $\mathcal{T}''$ (et, en particulier, de ne pas distinguer jusque dans leurs notations le vecteur y du covecteur ξ_y). Or le caractère semi-linéaire de l'isomorphisme en cas hermitien est à la base de certaines restrictions sur cette identification.

On tourne la difficulté si l'on prend pour covecteurs de $\mathcal{T}''$ non des fonctionnelles linéaires, mais des *fonctionnelles semi-linéaires* ξ : $\mathcal{T}' \to \mathbb{C}$, i.e. des applications telles que

$$\xi(\mathbf{x} + \mathbf{y}) = \xi(\mathbf{x}) + \xi(\mathbf{y})$$

et

$$\xi(c\mathbf{x}) = \bar{c}(\xi\mathbf{x})$$

quels que soient les vecteurs $\mathbf{x}$, $\mathbf{y} \in \mathcal{T}'$ et le nombre $c \in \mathbb{C}$. L'artifice est loin d'être universellement admis, mais il paraît qu'il a un grand avenir.

Un autre procédé consiste à munir $\mathcal{T}''$ formé de fonctionnelles linéaires $\xi : \mathcal{T}' \to \mathbb{C}$ de l'opération de multiplication de ξ par les nombres $c \in \mathbb{C}$: $\xi \mapsto c\xi$. On pose à cet effet

$$(c\xi)(\mathbf{x}) = \xi(\bar{c}\mathbf{x}) = \bar{c}(\xi(\mathbf{x}))$$

quels que soient la fonctionnelle $\xi \in \mathcal{T}''$, le nombre $c \in \mathbb{C}$ et le vecteur $\mathbf{x} \in \mathcal{T}'$.

Nous n'avons en fait que changé de place la conjugaison complexe « parasite » de sorte qu'elle a peut-être l'avantage de moins sauter aux yeux. Les deux méthodes ont chacune le bon et le mauvais côté. Ni l'une ni l'autre ne sont pas universelles, et elles exigent qu'on revoie les connaissances acquises antérieurement (disons, la théorie des tenseurs). Aussi nous en resterons à notre tactique ancienne sans soupirer après la perfection formelle que lesdits procédés aideraient à atteindre.

Quant à l'identification de vecteurs et de covecteurs, nous nous permettrons de la réaliser dans le cas hermitien sans oublier un instant l'éventualité de la conjugaison complexe fastidieuse.

* * *

L'identification *de facto* de vecteurs et de covecteurs des espaces euclidiens et unitaires autorise à en faire de même en ce qui concerne des objets de nature foncièrement différente dans les espaces vectoriels quelconques.

On voit par exemple qu'avec ladite identification de vecteurs et de covecteurs, l'annulateur S° d'un ensemble $S \subset \mathcal{V}^*$ quelconque coïncide avec le supplémentaire orthogonal $S^\perp$. En effet, le fait $\xi \in S^\circ$ signifie que $\xi(\mathbf{x}) = 0$ pour tout $\mathbf{x} \in S$. Si l'on identifie le covecteur ξ et un vecteur $\mathbf{y} \in \mathcal{V}$ tel qu'on ait $\xi(\mathbf{x}) = (\mathbf{x}, \mathbf{y})$ quel que soit $\mathbf{x} \in \mathcal{V}$, tout vecteur $\mathbf{x} \in S$ vérifie en particulier l'égalité $(\mathbf{x}, \mathbf{y}) = 0$ qui équivaut à $\mathbf{x} \in S^\perp$. □

Cela explique le parallélisme mentionné entre annulateurs et supplémentaires orthogonaux.

* * *

L'identification de vecteurs et de covecteurs en espace euclidien permet une simplification particulièrement notable de la théorie des tenseurs. En effet, on identifie des tenseurs de type (p, q) différents à une même somme $p + q$ parce qu'il ne dépend que de nous que telle variable tensorielle soit assimilée à un vecteur ou à un covecteur.

Soit par exemple un tenseur de type (2, 0), i.e. une fonctionnelle bilinéaire $A : \mathbf{x}, \mathbf{y} \mapsto A(\mathbf{x}, \mathbf{y})$. Si l'on prend $\mathbf{y}$ pour covecteur, on obtient une fonctionnelle bilinéaire de type (1,1), i.e. un opérateur linéaire $\mathbf{A} : \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$. Soyez attentifs, sinon vous risquez de vous embrouiller dans ces identifications : l'opérateur $\mathbf{A}$ transforme $\mathbf{x}$ en un vecteur $\mathbf{A}\mathbf{x}$ tel qu'étant considéré comme fonctionnelle sur des covecteurs il prenne sur ξ la valeur $\xi(\mathbf{A}\mathbf{x}) = A(\mathbf{x}, \mathbf{y})$, avec $\mathbf{y} = \xi$. Mais $\xi = \mathbf{y}$ signifie que $\xi(\mathbf{z}) = (\mathbf{z}, \mathbf{y})$ pour tout vecteur $\mathbf{z} \in \mathcal{V}$ et qu'on a en particulier $\xi(\mathbf{A}\mathbf{x}) = (\mathbf{A}\mathbf{x}, \mathbf{y})$. Ainsi,

$$A(\mathbf{x}, \mathbf{y}) = (\mathbf{A}\mathbf{x}, \mathbf{y}). \tag{1}$$

La formule (1) explicite la bijection entre les opérateurs linéaires $\mathbf{A} : \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ et les fonctionnelles bilinéaires $A : \mathbf{x}, \mathbf{y} \mapsto A(\mathbf{x}, \mathbf{y})$ sur l'espace euclidien $\mathcal{V}$. Si l'on faisait abstraction de la théorie générale, on l'accepterait en qualité de la définition de cette bijection. Dans ce cas, on doit certes établir que 1) la fonctionnelle A définie par (1) est bilinéaire pour tout opérateur linéaire $\mathbf{A}$ (on le vérifie automatiquement), 2) la correspondance « l'opérateur » $\mapsto$ « la fonctionnelle » est un homomorphisme d'espaces vectoriels associés (même remarque), 3) cet homomorphisme est injectif (on pose $\mathbf{y} = \mathbf{A}\mathbf{x}$ et on utilise la propriété de non-dégénérescence du produit scalaire) et 4) cet homomorphisme est un isomorphisme (la propriété découle de 3) parce que les deux espaces vectoriels sont de même dimension n^2).

La dernière approche reste valable dans le cas hermitien à la différence qu'on obtient des fonctionnelles sesquilinéaires et non des fonctionnelles bilinéaires. La nécessité d'en tenir constamment compte nous oblige à nous borner (dans cette leçon) aux espaces euclidiens. Un lecteur désireux de se placer dans le cas hermitien est déjà en mesure d'effectuer toutes les modifications qui s'imposent.

* * *

Un lecteur vigilant a visiblement remarqué un certain arbitraire laissé sur l'identification décrite de fonctionnelles bilinéaires et d'opérateurs linéaires. En effet, on prend pour covecteur la deuxième variable de A (x, y), mais cela pourrait être (dans l'espace euclidien) x et non y, auquel cas on obtiendrait en général un opérateur linéaire A* distinct de A tel que

$$A(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{A}^*\mathbf{y}). \tag{2}$$

La situation est la même (sinon plus mauvaise) en ce qui concerne les tenseurs d'autres types. Soit par exemple $T(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3; \xi_1)$ de type (3, 1). On assimile le vecteur $\mathbf{x}_3$ à un covecteur (qu'on désigne, disons, par ξ_2) et on identifie par là même le tenseur donné et $T(\mathbf{x}_1, \mathbf{x}_2; \xi_2, \xi_1)$ de type (2, 2). D'autre part, on peut prendre ξ_2 pour deuxième variable covectorielle, ce qui donne en général lieu à un autre tenseur de type (2, 2). Il y a plus. Si l'on fait jouer le rôle de covecteur non à $\mathbf{x}_3$, mais à $\mathbf{x}_2$, on construit un troisième tenseur de type (2, 2) autre que les deux précédents. On peut assimiler $\mathbf{x}_1$ à un covecteur et ξ_1 à un vecteur ! Un tenseur de type (3, 1) qui en provient diffère du tenseur primitif, et ainsi de suite.

On introduit pour s'y orienter une numérotation (ou au moins un ordonnancement) unique des variables vectorielles et des variables covectorielles qu'on écrit dans l'ordre donné. Ainsi, le fait de noter $T(\mathbf{x}_1, \mathbf{x}_2, \xi_3, \mathbf{x}_4)$ un tenseur de type (3, 1) signifie que si l'on assimile ξ_3 à un vecteur, le tenseur de type (4, 0) ainsi obtenu s'écrit $T(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4)$ et que si l'on identifie $\mathbf{x}_2$ (resp. $\mathbf{x}_4$) à un covecteur, on obtient le tenseur de type (2, 2) $T(\mathbf{x}_1, \xi_2, \xi_3, \mathbf{x}_4)$ (resp. $T(\mathbf{x}_1, \mathbf{x}_2, \xi_3, \xi_4)$).

Nous tenons à dire pour éviter tout malentendu que par exemple les symboles $T(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \xi_4)$ et $T(\mathbf{x}_1, \mathbf{x}_2, \xi_3, \mathbf{x}_4)$ désignent l'un et l'autre un tenseur de type (3, 1) à trois variables vectorielles et à une variable covectorielle. La seule chose qui les distingue entre eux est que si le premier s'obtient en qualifiant de covecteur la variable $\mathbf{x}_4$ d'un tenseur $T(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4)$ de type (4, 0), le second provient par identification analogue de la variable $\mathbf{x}_3$. Dans les espaces vectoriels arbitraires, cette différence n'a pas de sens.

* * *

Soit $\mathbf{e}_1, \ldots, \mathbf{e}_n$ une base quelconque de l'espace euclidien $\mathcal{T}^{\cdot}$. L'identification $\mathcal{T}^{\cdot} = \mathcal{T}^{\cdot\prime}$ fait que la base duale $\mathbf{e}^1, \ldots, \mathbf{e}^n$ est une base de $\mathcal{T}^{\cdot}$ qui diffère en général de $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Il y a entre les deux bases la relation

$$(\mathbf{e}_i, \mathbf{e}^j) = \delta_i^j, \quad i, j = 1, \ldots, n.$$

Si les formules

$$\mathbf{e}_i = g_{ij}\mathbf{e}^j, \; i, j = 1, \ldots, n,$$

régissent le passage de $\mathbf{e}^1, \ldots, \mathbf{e}^n$ à $\mathbf{e}_1, \ldots, \mathbf{e}_n$, alors

$$(\mathbf{e}_i, \mathbf{e}_j) = g_{ik}(\mathbf{e}^k, \mathbf{e}_j) = g_{ik}\delta_j^k = g_{ij},$$

où les nombres g_{ij} sont les *coefficients métriques* de $\mathbf{e}_1, \ldots, \mathbf{e}_n$ (voir I.14). Ils constituent une matrice non dégénérée dont l'inverse a pour éléments

$$g^{ij} = (\mathbf{e}^i, \mathbf{e}^j).$$

On passe à la base

$$\mathbf{e}_{i'} = c_{i'}^{i}\mathbf{e}_i$$

dont les coefficients métriques sont définis par

$$g_{i'j'} = c_{i'}^{i}c_{j'}^{j}g_{ij},$$

i.e. ils obéissent à la loi de transformation des tenseurs. Autrement dit, g_{ij} sont les coefficients d'un tenseur

$$G = g_{ij}\mathbf{e}^i \otimes \mathbf{e}^j = g^{ij}\mathbf{e}_i \otimes \mathbf{e}_j$$

appelé *tenseur métrique* de l'espace euclidien $\mathcal{T}^{\cdot}$. Sa valeur sur $\mathbf{x}, \mathbf{y} \in \mathcal{T}^{\cdot}$ est justement le produit scalaire de ces vecteurs:

$$G(\mathbf{x}, \mathbf{y}) = g_{ij}x^iy^j = (\mathbf{x}, \mathbf{y}).$$

Ainsi, il y a équivalence entre le terme « tenseur métrique » et le terme « produit scalaire ».

Soit $\mathbf{x}$ un vecteur arbitraire de $\mathcal{T}^{\cdot}$. Le produit tensoriel $\mathbf{x} \otimes G$ est par définition un tenseur de type (2, 1). On contracte (voir leçon 6) l'indice supérieur et le deuxième indice inférieur (le choix de l'indice inférieur est ici indifférent et on n'y a procédé que pour fixer les idées), il vient un tenseur de type (1, 0), i.e. un covecteur ξ. La valeur $\xi(\mathbf{y})$, $\mathbf{y}$ étant quelconque, est égale au produit tensoriel contracté $\operatorname{tr}(\xi \otimes \mathbf{y})$ et, partant, au résultat de la contraction totale de $\mathbf{x} \otimes G \otimes \mathbf{y}$, i.e. à $G(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{y})$, valeur du tenseur G sur les vecteurs $\mathbf{x}$ et $\mathbf{y}$ (cf. exemples de la leçon 6). Comme $\xi(\mathbf{y}) = (\mathbf{x}, \mathbf{y})$ signifie par définition l'identification du covecteur ξ et du vecteur $\mathbf{x}$, on vient de prouver que *le vecteur* $\mathbf{x}$ *considéré comme covecteur est le produit contracté du vecteur* $\mathbf{x}$ *et du tenseur* G. □

* * *

Dans la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$, les coordonnées du tenseur $\mathbf{x} \otimes G$ sont $g_{ij}x^k$ et celles du produit contracté de $\mathbf{x}$ et de G

$$x_i = g_{ij}x^j.$$

Les nombres $x_1, \ldots, x_n$ s'appellent les *coordonnées covariantes* du vecteur $\mathbf{x}$ dans la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Ce sont par définition les coordonnées du covecteur associé ξ dans la base duale $\mathbf{e}^1, \ldots, \mathbf{e}^n$ ou, ce qui revient au même, les coordonnées de $\mathbf{x}$ dans la base $\mathbf{e}^1, \ldots, \mathbf{e}^n$ dont les éléments sont identifiés aux vecteurs de $\mathcal{T}'$. Quant aux « vraies » coordonnées $x^1, \ldots, x^n$ de $\mathbf{x}$ dans la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$, elles sont dites *contravariantes*.

Le passage des x^i aux x_i et l'opération inverse sont appelés par certains auteurs l'*abaissement* et l'*élévation* de l'indice i respectivement.

L'ordonnancement décrit des variables des tenseurs oblige à en faire autant en ce qui concerne les indices inférieurs et supérieurs de leurs coordonnées (composantes). S'agissant d'un tenseur affecté de deux sortes d'indices, on laisse des blancs au-dessus des indices inférieurs et d'autres au-dessous des indices supérieurs. On rend cette écriture plus intuitive en remplaçant les blancs par les points.

Ainsi, les coordonnées de $T(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \xi_1)$ s'écrivent

$$T^{\cdot\,\cdot\,\cdot\,j_1}_{i_1 i_2 i_3 \cdot} = T_{i_1 i_2 i_3}{}^{j_1},$$

et celles de $T(\mathbf{x}_1, \mathbf{x}_2, \xi_1, \mathbf{x}_3)$ sont symbolisées par

$$T^{\cdot\,\cdot\,j_1 \cdot}_{i_1 i_2 \cdot i_3} = T_{i_1 i_2}{}^{j_1}{}_{i_3}.$$

Les coordonnées d'un opérateur linéaire sont en particulier désignées par $a_i{}^j$ ou $a^j{}_i$ selon que l'opérateur provient d'une fonctionnelle bilinéaire de coordonnées a_{ij} en assimilant à un covecteur la deuxième ou la première variable (formules (1) et (2) respectivement). Comme

$$a_{ij} = A(\mathbf{e}_i, \mathbf{e}_j), \quad a_i{}^j = (\mathbf{A}\mathbf{e}_i, \mathbf{e}^j), \quad a^j{}_i = (\mathbf{e}_i, \mathbf{A}^*\mathbf{e}^j),$$

on a

$$a_{ij} = g_{ik}a^k{}_j = g_{kj}a_i{}^k \tag{3}$$

et

$$a_i{}^j = g^{kj}a_{ik}, \quad a^i{}_j = g^{ik}a_{kj}. \tag{4}$$

Les nombres a_{ij}, $a_i{}^j$, $a^i{}_j$ (et $a^{ij} = g^{ik}g^{jl}a_{kl}$) peuvent être considérés comme diverses coordonnées d'un même être mathématique. A l'instar du corpuscule de la mécanique quantique, cet être présente deux aspects, un aspect « fonctionnel » et un aspect « opératoriel ». Les coordonnées a_{ij} sont dites purement covariantes, $a_i{}^j$ sont mixtes (elles sont covariantes en i et contravariantes en j), et ainsi de suite.

Selon les formules (3) et (4), ces coordonnées se déduisent les unes des autres en effectuant le produit tensoriel par les tenseurs « réciproques » g_{ij} et g^{ij} et en contractant les indices correspondants.

On élève et on abaisse de même les indices d'autres tenseurs. Par exemple,

$$T_{i_1 i_2}{}^{j_1 j_2} = g_{i_1 k_1} g^{j_1 l_1} T^{k_1}{}_{i_2 l_1}{}^{j_2}.$$

Si la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ est orthonormée, la base duale $\mathbf{e}^1, \ldots, \mathbf{e}^n$ lui est égale, si bien que toutes les formules d'abaissement et d'élévation des indices deviennent les égalités entre les coordonnées associées de mêmes indices (la position des indices étant indifférente). Ainsi,

$$a_{ij} = a_i{}^j = a^j{}_i \tag{5}$$

pour les fonctionnelles bilinéaires et

$$x_i = x^i$$

pour les vecteurs. Ce n'est donc pas pour rien que dès le premier semestre nous avons indicé inférieurement les coordonnées des vecteurs dans une base orthonormée.

D'après la première égalité (5), la fonctionnelle bilinéaire A et l'opérateur linéaire $\mathbf{A}$ associé par la formule (1) ont même matrice dans toute base orthonormée.

Quant à l'opérateur $\mathbf{A}^*$ défini par (2), sa matrice (dans une base orthonormée) coïncide avec la transposée de la matrice de $\mathbf{A}$.

Dans les pages suivantes, nous identifierons toujours les fonctionnelles bilinéaires et les opérateurs linéaires à l'aide de la formule (1), si bien que nous emploierons, pour désigner les éléments de la matrice représentative d'un opérateur linéaire, la notation usuelle a^i_j au lieu de la notation plus précise $a^i{}_j$.

LEÇON 19

Opérateurs adjoints. — Opérateurs auto-adjoints. — Opérateurs antisymétriques et antihermitiens. — Une analogie entre les opérateurs hermitiens et les nombres réels. — Propriétés spectrales des opérateurs auto-adjoints. — Propriété des opérateurs auto-adjoints d'être orthogonalement diagonalisables.

Conformément aux formules (1) et (2) de la leçon 18, on associe à chaque opérateur linéaire **A** d'un espace euclidien (ou hermitien) $\mathcal{V}$ dans lui-même une fonctionnelle bilinéaire A et à celle-ci l'opérateur linéaire **A***.

Définition 1. L'opérateur **A*** est dit *adjoint* de **A**. Il est bien défini par la relation

$$(\mathbf{Ax}, \mathbf{y}) = (\mathbf{x}, \mathbf{A}^*\mathbf{y}) \tag{1}$$

valable pour tout **x** et tout **y** de $\mathcal{V}$.

La définition a un sens pour $\mathcal{V}$ euclidien ou hermitien. Il y a cependant une différence. Dans le premier cas, **A*** est en fait l'opérateur $\mathbf{A}' : \mathcal{V}' \to \mathcal{V}'$ considéré comme agissant sur $\mathcal{V}$ par suite de l'identification de $\mathcal{V}$ et de $\mathcal{V}'$, tandis que dans le second, l'opérateur **A*** diffère de **A**′ par la conjugaison complexe même après identification des vecteurs et des covecteurs.

Dans une base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ quelconque de $\mathcal{V}$ euclidien, les éléments a^{*i}_j de la matrice de **A*** sont liés aux éléments a^i_j de la matrice de **A** par la formule

$$a^{*i}_j = g^{ik} g_{jl} a^l_k.$$

S'agissant de $\mathbf{e}_1, \ldots, \mathbf{e}_n$ orthonormée, elle s'écrit

$$a^{*i}_j = a^j_i.$$

La formule correspondante (dans une base orthonormée) est pour $\mathcal{V}$ hermitien :

$$a^{*i}_j = \bar{a}^j_i.$$

Ainsi, *l'opérateur* **A*** *sur un espace euclidien* $\mathcal{V}$ (*resp. hermitien*) *est l'adjoint de l'opérateur* **A** *si et seulement si sa matrice dans une* (*donc dans toute*) *base orthonormée est égale à la transposée* (*resp. la transposée de la conjuguée*) *de la matrice de* **A**. □

Dans le cas euclidien, on le démontre directement si l'on se rappelle que **A** et **A**′ sont représentés dans des bases duales par les

matrices transposées (voir leçon 14) et que la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ est orthonormée si et seulement si elle coïncide avec la duale $\mathbf{e}^1, \ldots, \mathbf{e}^n$ considérée comme base dans $\mathcal{T}^*$.

Il est naturel que $\mathbf{A}^*$ et $\mathbf{A}'$ jouissent des propriétés complètement analogues. Par exemple, $\mathbf{A}^{**} = \mathbf{A}$ et $(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*$. La seule différence qui compte est redevable une fois de plus à la multiplication par les nombres en espace hermitien. En effet, la formule $(c\mathbf{A})^* = c\mathbf{A}^*$ (cas euclidien) devient $(c\mathbf{A})^* = \bar{c}\mathbf{A}^*$ en espace hermitien, i.e. on voit apparaître la conjugaison complexe « parasite ».

* * *

Nous allons donner une définition qui s'appuie sur la propriété de $\mathbf{A}$ et $\mathbf{A}^*$ d'opérer dans un même espace (elle n'a donc pas de sens pour $\mathbf{A}'$).

Définition 2. Un opérateur $\mathbf{A}$ de $\mathcal{T}^*$ euclidien ou hermitien dans lui-même est dit *auto-adjoint* si $\mathbf{A} = \mathbf{A}^*$, i.e. si l'on a pour deux vecteurs $\mathbf{x}, \mathbf{y} \in \mathcal{T}^*$ quelconques

$$(\mathbf{Ax}, \mathbf{y}) = (\mathbf{x}, \mathbf{Ay}).$$

Selon qu'il s'agit d'un espace euclidien ou d'un espace hermitien, les opérateurs auto-adjoints sont appelés *symétriques* ou *hermitiens*.

Il est clair qu'*un opérateur* $\mathbf{A}$ *sur un espace euclidien* (resp. *hermitien*) *est symétrique* (resp. *hermitien*) *si et seulement si la fonctionnelle bilinéaire* (resp. *sesquilinéaire*) *associée* A *est symétrique* (resp. *hermitienne*).

On a par exemple dans l'espace hermitien

$$A(\mathbf{y}, \mathbf{x}) = (\mathbf{Ay}, \mathbf{x}) = \overline{(\mathbf{x}, \mathbf{Ay})} = \overline{(\mathbf{Ax}, \mathbf{y})} = \overline{A(\mathbf{x}, \mathbf{y})}. \quad \square$$

La somme d'opérateurs auto-adjoints est évidemment un opérateur auto-adjoint, ainsi que le produit d'un opérateur auto-adjoint par un nombre réel. Autrement dit, *les opérateurs auto-adjoints constituent un espace vectoriel sur* $\mathbb{R}$ (qui est un sous-espace de $\mathrm{Op}(\mathcal{T}^*)$ pour $\mathcal{T}^*$ euclidien).

On note que tout produit de deux opérateurs auto-adjoints n'est pas un opérateur auto-adjoint. On dit plus précisément que *le produit* $\mathbf{AB}$ *des opérateurs auto-adjoints* $\mathbf{A}$ *et* $\mathbf{B}$ *est un opérateur auto-adjoint si et seulement si* $\mathbf{A}$ *et* $\mathbf{B}$ *sont commutables, i.e. si* $\mathbf{AB} = \mathbf{BA}$.

En effet, $(\mathbf{AB})^* = (\mathbf{BA})^* = \mathbf{A}^*\mathbf{B}^* = \mathbf{AB}$ si $\mathbf{AB} = \mathbf{BA}$. Inversement, $\mathbf{BA} = \mathbf{B}^*\mathbf{A}^* = (\mathbf{AB})^* = \mathbf{AB}$ quand $(\mathbf{AB})^* = \mathbf{AB}$. $\square$

Une matrice carrée $A = (a_i^j)$ à éléments complexes est dite *hermitienne* si elle est égale à la matrice transposée de sa conjuguée, i.e. si

$$a_i^j = \bar{a}_j^i \text{ quels que soient } i, j = 1, \ldots, n.$$

Il est clair qu'*une condition nécessaire et suffisante pour qu'un opérateur* **A** *dans un espace euclidien* (resp. *hermitien*) *soit symétrique* (resp. *hermitien*) *est que la matrice associée à* **A** *dans une* (*donc dans toute*) *base orthonormée soit symétrique* (resp. *hermitienne*). □

* * *

Définition 3. Un opérateur **A** sur un espace euclidien $\mathscr{V}$ est dit *antisymétrique* si $\mathbf{A}^* = -\mathbf{A}$, i.e. si

$$(\mathbf{Ax}, \mathbf{y}) + (\mathbf{x}, \mathbf{Ay}) = 0$$

quels que soient $\mathbf{x}, \mathbf{y} \in \mathscr{V}$.

On dit de même d'un opérateur **A** d'un espace unitaire qu'il est *antihermitien* si $\mathbf{A}^* = -\mathbf{A}$.

Les opérateurs antisymétriques forment une classe autonome d'opérateurs linéaires. On leur attache les fonctionnelles bilinéaires antisymétriques et on les caractérise en coordonnées par l'antisymétrie des matrices associées dans les bases orthonormées. Ces opérateurs constituent un sous-espace de Op ($\mathscr{V}$), l'espace Op ($\mathscr{V}$) des opérateurs linéaires étant somme directe du sous-espace des opérateurs symétriques et du sous-espace des opérateurs antisymétriques (cf. leçon 11, prop. 1), i.e. chaque opérateur linéaire **A** est répresenté de façon unique par la somme

$$\mathbf{A} = \mathbf{A}_{\text{sym}} + \mathbf{A}_{\text{anti}} \tag{2}$$

d'un opérateur symétrique $\mathbf{A}_{\text{sym}}$ et d'un opérateur antisymétrique $\mathbf{A}_{\text{anti}}$. Cela étant, on a

$$\mathbf{A}_{\text{sym}} = \frac{\mathbf{A}+\mathbf{A}^*}{2}, \quad \mathbf{A}_{\text{anti}} = \frac{\mathbf{A}-\mathbf{A}^*}{2}.$$

Le cas d'opérateurs hermitiens s'avère tout autre parce que les opérateurs antihermitiens se réduisent de façon triviale aux opérateurs hermitiens (propriété spéciale aux espaces unitaires). Plus précisément, la relation $(i\mathbf{A})^* = \bar{i}\mathbf{A}^* = -i\mathbf{A}^*$ entraîne de suite que *pour qu'un opérateur soit antihermitien il faut et il suffit qu'il s'écrive* $i\mathbf{A}$, **A** *étant un opérateur hermitien.* □

* * *

D'autre part, les opérateurs sur les espaces unitaires se décomposent évidemment en une somme directe analogue à (2). Par conséquent, *tout opérateur* **A** *sur un espace hermitien* $\mathscr{V}$ *est représentable de façon unique par la somme*

$$\mathbf{A} = \mathbf{B} + i\mathbf{C},$$

B *et* **C** *étant des opérateurs hermitiens.* Cela signifie qu'*étant donné un espace hermitien* $\mathscr{V}$ *quelconque, l'espace vectoriel* Op ($\mathscr{V}$) *est*

muni de la structure naturelle de ($\mathbb{C}$, $\mathbb{R}$)*-espace vectoriel, le sous-espace réel associé étant l'espace des opérateurs hermitiens* (voir I. 19, déf. 1). □

On voit que les opérateurs hermitiens sont analogues en un sens aux nombres réels. On retrouve cette analogie dans d'autres questions.

Selon la définition 2 de la leçon 11 et la relation établie plus haut entre les fonctionnelles bilinéaires symétriques et les opérateurs linéaires symétriques, *chaque fonctionnelle quadratique sur un espace euclidien s'écrit de façon unique sous forme* (Ax, x), A *étant un opérateur linéaire symétrique.* S'agissant d'opérateurs linéaires non symétriques, ces fonctionnelles ne donnent lieu à aucun résultat nouveau parce que $(\mathbf{Ax}, \mathbf{x}) = 0$ pour tous les $\mathbf{x} \in \mathcal{V}$ si (et seulement si) A est antisymétrique.

Il n'en est rien pour ce qui est des espaces unitaires. On s'y attend d'ailleurs du moment qu'aucune fonctionnelle (Ax, x), $\mathbf{A} \neq 0$, d'un espace hermitien n'est une fonctionnelle quadratique au sens de la définition mentionnée. Pour quelle raison ses propriétés rappelleraient-elles celles des fonctionnelles quadratiques ? !

Dans le cas euclidien, la fonctionnelle (Ax, x) peut être identiquement nulle sans que $\mathbf{A} = 0$. La chose est impossible pour les espaces hermitiens.

Proposition 1. *Si l'opérateur linéaire* A *d'un espace hermitien* $\mathcal{V}$ *dans lui-même est tel que*

$$(\mathbf{Ax}, \mathbf{x}) = 0 \tag{3}$$

pour tout vecteur $\mathbf{x} \in \mathcal{V}$, *alors* $\mathbf{A} = 0$.

Démonstration. On a pour tout x et tout y de $\mathcal{V}$

$$(\mathbf{A}(\mathbf{x} + \mathbf{y}), \mathbf{x} + \mathbf{y}) = (\mathbf{Ax}, \mathbf{x}) + (\mathbf{Ax}, \mathbf{y}) + (\mathbf{Ay}, \mathbf{x}) + (\mathbf{Ay}, \mathbf{y}),$$
$$(\mathbf{A}(\mathbf{x} + i\mathbf{y}), \mathbf{x} + i\mathbf{y}) = (\mathbf{Ax}, \mathbf{x}) + (\mathbf{Ax}, i\mathbf{y}) + (i\mathbf{Ay}, \mathbf{x}) + (i\mathbf{Ay}, i\mathbf{y})$$

et

$$(\mathbf{Ax}, i\mathbf{y}) = -i(\mathbf{Ax}, \mathbf{y}), \quad (i\mathbf{Ay}, \mathbf{x}) = i(\mathbf{Ax}, \mathbf{y}),$$

si bien que la condition (3) entraîne

$$(\mathbf{Ax}, \mathbf{y}) + (\mathbf{Ay}, \mathbf{x}) = 0,$$
$$(\mathbf{Ax}, \mathbf{y}) - (\mathbf{Ay}, \mathbf{x}) = 0,$$

et, partant, $(\mathbf{Ax}, \mathbf{y}) = 0$. On pose $\mathbf{y} = \mathbf{Ax}$, il vient $(\mathbf{Ax}, \mathbf{Ax}) = 0$, donc $\mathbf{Ax} = 0$ quel que soit $\mathbf{x} \in \mathcal{V}$. □

Proposition 2 (critère d'opérateur hermitien). *Un opérateur linéaire* A *sur un espace hermitien* $\mathcal{V}$ *est hermitien si et seulement si le nombre* (Ax, x) *est réel quel que soit le vecteur* $\mathbf{x} \in \mathcal{V}$.

Démonstration. Si A est un opérateur hermitien, alors tout vecteur $\mathbf{x} \in \mathcal{V}$ vérifie les égalités

$$(\mathbf{Ax}, \mathbf{x}) = (\mathbf{x}, \mathbf{Ax}) = \overline{(\mathbf{Ax}, \mathbf{x})},$$

donc $(\mathbf{Ax}, \mathbf{x})$ est réel. Inversement, soit $(\mathbf{Ax}, \mathbf{x}) = \overline{(\mathbf{Ax}, \mathbf{x})}$, auquel cas

$$((\mathbf{A}-\mathbf{A}^*)\,\mathbf{x}, \mathbf{x}) = (\mathbf{Ax}, \mathbf{x}) - (\mathbf{A}^*\mathbf{x}, \mathbf{x}) = (\mathbf{Ax}, \mathbf{x}) - (\mathbf{x}, \mathbf{Ax}) =$$
$$= (\mathbf{Ax}, \mathbf{x}) - \overline{(\mathbf{Ax}, \mathbf{x})} = 0,$$

ce qui entraîne $\mathbf{A} - \mathbf{A}^* = 0$ par la proposition 1. □

* * *

Les propositions suivantes s'appliquent au cas euclidien comme au cas hermitien (encore que la démonstration soit en général différente).

Proposition 3 (sur les racines caractéristiques réelles). *Tout opérateur auto-adjoint a ses racines caractéristiques réelles.*

Démonstration. Soient $\mathbf{A}$ un opérateur auto-adjoint dans $\mathcal{V}$ euclidien ou hermitien et λ sa racine caractéristique quelconque.

Si $\mathcal{V}$ est unitaire (i.e. $\mathbf{A}$ est hermitien), le nombre λ constitue une valeur propre de $\mathbf{A}$, i.e. il existe un vecteur $\mathbf{x}_0 \neq 0$ tel que $\mathbf{Ax}_0 = \lambda\mathbf{x}_0$, auquel cas $(\mathbf{Ax}_0, \mathbf{x}_0) = (\lambda\mathbf{x}_0, \mathbf{x}_0) = \lambda\,(\mathbf{x}_0, \mathbf{x}_0)$, donc

$$\lambda = \frac{(\mathbf{Ax}_0, \mathbf{x}_0)}{(\mathbf{x}_0, \mathbf{x}_0)}.$$

On a l'affirmation voulue si l'on dit que le second membre de la formule ci-dessus est réel (par la proposition 2), ce qui implique la même propriété de λ.

On se place maintenant dans $\mathcal{V}$ euclidien. On raisonne par l'absurde et on suppose que $\lambda = \alpha + i\beta$, avec $\beta \neq 0$. Il est connu dès la leçon 16 qu'il existe pour $\mathbf{A}$ dans l'espace $\mathcal{V}$ un sous-espace invariant $\mathcal{P}$ de dimension 2 et une base $\mathbf{x}$, $\mathbf{y}$ de $\mathcal{P}$ tels que

$$\mathbf{Ax} = \alpha\mathbf{x} - \beta\mathbf{y},$$
$$\mathbf{Ay} = \beta\mathbf{x} + \alpha\mathbf{y}.$$

Aussi

$$(\mathbf{Ax}, \mathbf{y}) = (\alpha\mathbf{x} - \beta\mathbf{y}, \mathbf{y}) = \alpha\,(\mathbf{x}, \mathbf{y}) - \beta\,(\mathbf{y}, \mathbf{y})$$

et

$$(\mathbf{x}, \mathbf{Ay}) = (\mathbf{x}, \beta\mathbf{x} + \alpha\mathbf{y}) = \beta\,(\mathbf{x}, \mathbf{x}) + \alpha\,(\mathbf{x}, \mathbf{y}).$$

L'opérateur $\mathbf{A}$ étant auto-adjoint (symétrique), on a $(\mathbf{Ax}, \mathbf{y}) = (\mathbf{x}, \mathbf{Ay})$, d'où

$$\beta\,[(\mathbf{x}, \mathbf{x}) + (\mathbf{y}, \mathbf{y})] = 0,$$

contradiction (car $(\mathbf{x}, \mathbf{x}) > 0$, $(\mathbf{y}, \mathbf{y}) > 0$ et $\beta \neq 0$ par hypothèse) qui démontre la propriété « $\lambda \in \mathbb{R}$ ». □

Proposition 4 (sur les vecteurs propres orthogonaux). *Etant donné un opérateur auto-adjoint* $\mathbf{A}$, *ses vecteurs propres quelconques* $\mathbf{x}$, $\mathbf{y}$ *associés aux valeurs propres* λ, μ *distinctes sont orthogonaux.*

Démonstration. On a

$$(\mathbf{Ax}, \mathbf{y}) = (\lambda\mathbf{x}, \mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y}),$$
$$(\mathbf{x}, \mathbf{Ay}) = (\mathbf{x}, \mu\mathbf{y}) = \mu(\mathbf{x}, \mathbf{y})$$

(la dernière égalité reste en vigueur dans le cas hermitien du moment que μ est réel par la proposition 3). La propriété de **A** d'être auto-adjoint implique donc

$$\lambda(\mathbf{x}, \mathbf{y}) = \mu(\mathbf{x}, \mathbf{y}),$$

égalité qui n'est possible pour λ et μ distinctes que si $(\mathbf{x}, \mathbf{y}) = 0$. □

Proposition 5 (sur le supplémentaire orthogonal). *Etant donné un opérateur auto-adjoint* **A** *arbitraire, le supplémentaire orthogonal* $\mathcal{P}^{\perp}$ *d'un sous-espace invariant* $\mathcal{P}$ *quelconque est un sous-espace invariant.*

Démonstration. Si $\mathbf{x} \in \mathcal{P}^{\perp}$, alors $(\mathbf{x}, \mathbf{y}) = 0$ pour tous les vecteurs **y** de $\mathcal{P}$, donc $(\mathbf{Ax}, \mathbf{y}) = (\mathbf{x}, \mathbf{Ay}) = 0$ car $\mathbf{Ay} \in \mathcal{P}$ par hypothèse. Par conséquent, $\mathbf{Ax} \in \mathcal{P}^{\perp}$. □

Proposition 6 (sur les multiplicités). *La multiplicité géométrique* p_{λ_0} *d'une valeur propre* λ_0 *quelconque d'un opérateur auto-adjoint* **A** *est égale à sa multiplicité algébrique* n_{λ_0} :

$$p_{\lambda_0} = n_{\lambda_0}.$$

Démonstration. Soient $\mathcal{P}_{\lambda_0}$ le sous-espace propre associé à la valeur propre λ_0 et $\mathbf{e}_1, \ldots, \mathbf{e}_n$ une base orthonormée de $\mathcal{V}$ telle que $\mathcal{P}_{\lambda_0} = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$ (i.e. telle que $\mathcal{P}_{\lambda_0}^{\perp} = [\mathbf{e}_{p+1}, \ldots, \mathbf{e}_n]$), avec $p = p_{\lambda_0}$. D'après la proposition 5,

$$\mathcal{P}_{\lambda_0} \oplus \mathcal{P}_{\lambda_0}^{\perp} = \mathcal{V},$$

si bien que la matrice de **A** dans cette base s'écrit

$$\left(\begin{array}{ccc|c} \lambda_0 & & & \\ & \ddots & & 0 \\ & & \lambda_0 & \\ \hline & 0 & & B \end{array}\right),$$

B étant la matrice de l'opérateur $\mathbf{B} = \mathbf{A}|_{\mathcal{P}_{\lambda_0}^{\perp}}$. Par conséquent, $f_{\mathbf{A}}(\lambda) = (\lambda_0 - \lambda)^p f_{\mathbf{B}}(\lambda)$. Aussi $f_{\mathbf{B}}(\lambda_0) = 0$ pour $p_{\lambda_0} < n$, et λ_0 est valeur propre de **B**. Tout vecteur propre associé de $\mathcal{P}_{\lambda_0}^{\perp}$ est vecteur propre de **A** relatif à la valeur propre λ_0, ce qui est impossible parce que tous les vecteurs propres de **A** associés à λ_0 sont dans $\mathcal{P}_{\lambda_0}$. Par conséquent, $p_{\lambda_0} \geqslant n_{\lambda_0}$ et, partant, $p_{\lambda_0} = n_{\lambda_0}$ (étant donné qu'on a nécessairement $p_{\lambda_0} \leqslant n_{\lambda_0}$; voir leçon 15). □

Remarque. La seule propriété utilisée de l'opérateur auto-adjoint est que le supplémentaire orthogonal de chaque sous-espace propre constitue un sous-espace invariant (i.e. nous n'avons eu besoin que d'une partie de l'affirmation 5). Aussi *la proposition* 6 *est vraie pour tout opérateur tel que le supplémentaire orthogonal de son sous-espace propre quelconque soit invariant.* □

* * *

Aux termes du théorème 1 de la leçon 16, la proposition 6 (jointe à la proposition 3 dans le cas euclidien) implique que **A** est diagonalisable, i.e.

$$\mathcal{V} = \mathcal{P}_{\lambda_1} \oplus \dots \oplus \mathcal{P}_{\lambda_m},$$

$\lambda_1, \dots, \lambda_m$ étant toutes les valeurs propres de **A**. Si l'on prend dans chaque sous-espace $\mathcal{P}_{\lambda_i}$ une base orthonormée, on obtient une base orthonormée de $\mathcal{V}$ (proposition 4) dans laquelle la matrice associée à **A** est diagonale.

Définition 4. Un opérateur **A** dans un espace euclidien ou hermitien $\mathcal{V}$ est dit *orthogonalement diagonalisable* s'il existe une base orthonormée de $\mathcal{V}$ dans laquelle la matrice de **A** est diagonale (i.e. une base constituée de vecteurs propres de **A**).

Ainsi, on vient de démontrer le

Théorème 1. *Tout opérateur auto-adjoint d'un espace euclidien ou hermitien est orthogonalement diagonalisable.* □

LEÇON 20

Réduction canonique des formes quadratiques par une transformation orthogonale de variables. — Hypersurfaces du second degré dans l'espace ponctuel euclidien. — Propriété de maximum-minimum des valeurs propres des opérateurs auto-adjoints. — Opérateurs orthogonalement diagonalisables.

Le théorème 1 de la leçon précédente dit que tout opérateur auto-adjoint d'un espace euclidien quelconque est orthogonalement diagonalisable. Voyons ce que sera cette affirmation en termes de formes bilinéaires symétriques (ou, ce qui revient au même, en termes de formes quadratiques).

Soit

$$Q(x_1, \ldots, x_n) = \sum_{i,\, j=1}^{n} q_{ij} x_i x_j \tag{1}$$

une forme quadratique quelconque de n variables $x_1, \ldots, x_n$ à coefficients réels q_{ij}, $i, j = 1, \ldots, n$.

On prend dans l'espace euclidien $\mathcal{T}$ de dimension n une base orthonormée $\mathbf{e}_1, \ldots, \mathbf{e}_n$ et on considère dans $\mathcal{T}$ une fonctionnelle quadratique Q représentée dans cette base par la forme $Q(x_1, \ldots, x_n)$, donc l'opérateur linéaire symétrique associé $\mathbf{Q}: \mathcal{T} \to \mathcal{T}$ (i.e. $\mathbf{Q}$ tel que $Q(\mathbf{x}) = (\mathbf{Qx}, \mathbf{x})$ pour tout vecteur $\mathbf{x}$ de $\mathcal{T}$). Selon le même théorème, il existe une base orthonormée $\mathbf{f}_1, \ldots, \mathbf{f}_n$ de $\mathcal{T}$ dans laquelle la matrice de $\mathbf{Q}$ est diagonale d'éléments diagonaux $\lambda_1, \ldots, \lambda_n$. Autrement dit, tout vecteur $\mathbf{x}$ de $\mathcal{T}$ satisfait à l'égalité

$$Q(\mathbf{x}) = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2,$$

où

$$\begin{aligned} y_1 &= c_{11}x_1 + \ldots + c_{1n}x_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n &= c_{n1}x_1 + \ldots + c_{nn}x_n \end{aligned} \tag{2}$$

sont les coordonnées de $\mathbf{x}$ par rapport à $\mathbf{f}_1, \ldots, \mathbf{f}_n$. Comme $\mathbf{e}_1, \ldots, \mathbf{e}_n$ et $\mathbf{f}_1, \ldots, \mathbf{f}_n$ sont orthonormées, la transformation (2) est orthogonale, i.e. la matrice C de ses coefficients est orthogonale (voir I.14). Cela démontre le

Théorème 1. *Toute forme quadratique* (1) *est ramenée par transformation orthogonale de variables à*

$$\lambda_1 y_1^2 + \ldots + \lambda_n y_n^2. \tag{3}$$

Cela étant, les coefficients $\lambda_1, \ldots, \lambda_n$ *sont racines de l'équation*

$$\det(Q - \lambda E) = 0,$$

si bien qu'ils sont définis de façon unique (à l'ordre près). □

La seule différence avec le théorème de Lagrange (sensiblement plus simple) de la leçon 11 est formelle : on réduit à la forme canonique (3) par la transformation orthogonale de variables (2) et non par une transformation arbitraire. D'où l'unicité de (3).

* * *

On se souvient que le théorème de Lagrange a permis d'établir une classification des hypersurfaces du second degré de l'espace affine n-dimensionnel (voir leçon 13). Avec le théorème 1, on procède de même en ce qui concerne l'espace ponctuel euclidien (réel-complexe) de dimension n. En effet, on reproduit exactement la démonstration du théorème 5 de la leçon 13 à la lumière du théorème 1 (au lieu du résultat de Lagrange), et on est conduit de suite au

Théorème 2 (réduction des équations des hypersurfaces du second degré de l'espace euclidien de dimension n à la forme canonique). *Pour toute hypersurface du second degré de l'espace euclidien réel-complexe de dimension* n $(n \geqslant 1)$, *il existe un système de coordonnées rectangulaires* $x_1, \ldots, x_n$ *en lesquelles l'équation de l'hypersurface est, ou bien*

$$\text{(I)} \qquad \lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 = \varepsilon,$$

avec $1 \leqslant r \leqslant n$, $\varepsilon = 0, 1$, *ou bien (pour* $n > 1$ *seulement)*

$$\text{(II)} \qquad \lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 = 2x_{r+1},$$

avec $1 \leqslant r \leqslant n - 1$, *et on a* $\lambda_1 \neq 0, \ldots, \lambda_r \neq 0$ *dans les deux cas.* □

Si l'on se propose de fixer de façon unique les coefficients $\lambda_1, \ldots, \lambda_r$ (proportionnels aux racines non nulles comptées avec leur multiplicité du polynôme caractéristique associé), on doit tout d'abord les ordonner de manière raisonnable (i.e. échanger convenablement les coordonnées $x^1, \ldots, x^n$). Nous exigerons qu'on dénombre en premier lieu les coefficients positifs, puis ceux négatifs. Les éléments de chaque groupe seront rangés dans l'ordre de croissance des valeurs absolues. Si p, $0 \leqslant p \leqslant r$, est le nombre de coefficients positifs, on estime donc que

$$0 < \lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_p$$

et

$$0 < |\lambda_{p+1}| \leqslant |\lambda_{p+2}| \leqslant \ldots \leqslant |\lambda_r|.$$

Si l'on multiplie par -1 dans le cas (I), $\varepsilon = 0$, on a

$$(4) \qquad 0 \leqslant p \leqslant \left[\frac{r}{2}\right].$$

On aboutit à ce résultat dans le cas (II) en changeant, si besoin est, le signe de x_{r+1}. Pour des raisons d'homogénéité, nous admettrons donc que $\varepsilon = -1$ pour (I), et nous réaliserons de sorte la condition (4).

Si on est dans le cas (I), $\varepsilon = 0$, on admet que

$$|\lambda_1| + \ldots + |\lambda_r| = 1.$$

Nous dirons des équations (I) et (II) vérifiant ces conditions qu'elles sont les *équations canoniques euclidiennes* des hypersurfaces du second degré.

Quand $n = 2$ et $n = 3$, ce sont évidemment les équations canoniques (aux notations près) des courbes du second degré dans le plan et des surfaces du second degré dans l'espace de dimension 3 dont la liste a été dressée dans I.22 et I.23.

Avec la méthode de réduction canonique du théorème 2 (i.e. avec la méthode de la leçon 13 à la différence que le résultat de Lagrange fait ici place au théorème 1), on a, on le constate facilement, une même équation canonique (éventuellement rapportée à des systèmes de coordonnées différents). Cela ne prouve certes pas que l'hypersurface ne peut être décrite dans certaines coordonnées par une autre équation canonique, mais l'on a néanmoins le

Théorème 3 (classification des hypersurfaces du second degré dans le ($\mathbb{C}$, $\mathbb{R}$)-espace euclidien de dimension n). *Deux hypersurfaces du second degré du ($\mathbb{C}$, $\mathbb{R}$)-espace euclidien de dimension n sont équivalentes euclidiennes si et seulement si elles ont une même équation canonique.*

On se souvient, pour démontrer le théorème, du cas de courbes du second degré dans le plan (voir I.22). Le procédé consiste à caractériser géométriquement les coefficients $\lambda_1, \ldots, \lambda_r$ en faisant abstraction des coordonnées. Voyons quelle est l'idée de la méthode générale. Soit, dans le plan, l'ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

$a \geqslant b$ $\left(\text{auquel cas } \lambda_1 = \frac{1}{a^2},\ \lambda_2 = \frac{1}{b^2}\right)$.

Le premier membre de l'équation de l'ellipse est une forme quadratique des coordonnées x, y des points du plan. Si $x^2 + y^2 = 1$ (« circonférence unité »), on voit aisément que les valeurs extrêmes de la forme sont $\lambda_2 = \frac{1}{b^2}$ (maximum) et $\lambda_1 = \frac{1}{a^2}$ (minimum).

Soit l'ellipsoïde

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \quad a \geqslant b \geqslant c.$$

Par analogie avec le cas précédent, le coefficient $\frac{1}{c^2}$ est égal à la plus grande valeur que la forme quadratique $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}$ prend sur la « sphère unité » $x^2 + y^2 + z^2 = 1$, et le coefficient $\frac{1}{a^2}$ à la valeur minimale. La caractérisation géométrique du coefficient « moyen » $\frac{1}{b^2}$ est plus difficile. On coupe à cet effet l'ellipsoïde par divers plans passant par son centre. Les sections sont des ellipses pour lesquelles on définit les coefficients λ_1, $\lambda_2 \geqslant \lambda_1$ correspondants. Ces coefficients dépendent bien sûr du plan, et on vérifie de suite que *la plus petite valeur possible du plus grand coefficient* λ_2 est justement $\frac{1}{b^2}$.

Il se trouve que la caractérisation décrite des $\lambda_1, \ldots, \lambda_2$ a également lieu dans le cas général. Le support en est un résultat sur les valeurs propres des opérateurs. Nous nous bornerons à ce résultat car le temps nous manque pour passer aux coefficients des équations des hypersurfaces, ce passage étant d'ailleurs absolument trivial.

* * *

Ainsi, on se place une fois de plus dans l'espace vectoriel euclidien $\mathcal{V}$, et soit A un opérateur symétrique défini sur $\mathcal{V}$. D'ailleurs, les formulations et les démonstrations ci-dessous restent entières pour $\mathcal{V}$ unitaire et A hermitien.

Dans les deux cas, l'opérateur A a toutes ses valeurs propres (= racines caractéristiques) réelles (leçon 19, prop. 3). On prend chacune des valeurs propres un nombre de fois égal à sa multiplicité (elles sont donc exactement n), et on les numérote dans l'ordre de décroissance :

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n.$$

Notre but consiste à les caractériser de manière « géométrique » directe.

Soient $\mathcal{P}$ un sous-espace quelconque de $\mathcal{V}$ et $S = S(\mathcal{P})$ sa partie (la « sphère unité ») formée de tous les vecteurs $\mathbf{x} \in \mathcal{P}$ pour lesquels $(\mathbf{x}, \mathbf{x}) = 1$.

Le nombre $(\mathbf{Ax}, \mathbf{x})$ étant réel pour tout $\mathbf{x} \in S$ (c'est évident pour $\mathcal{V}$ euclidien, et, dans le cas hermitien, la propriété est garantie par la proposition 2 de la leçon précédente), on définit le nombre

$$\alpha(\mathcal{P}) = \sup \{(\mathbf{Ax}, \mathbf{x}) ; \mathbf{x} \in \mathcal{P}, (\mathbf{x}, \mathbf{x}) = 1\}$$

(d'ailleurs, la sphère $S(\mathcal{P})$ est compacte, ce qui permet de remplacer sup par max).

Proposition 1. *On a pour tout* $q = 1, \ldots, n$ *l'égalité*

$$\lambda_q = \inf \{\alpha(\mathcal{P});\ \dim \mathcal{P} = n - q + 1\},$$

inf *étant étendu sur tous les sous-espaces* $\mathcal{P} \subset \mathcal{V}$ *de dimension* $n - q + 1$.

Démonstration. Aux termes du théorème 1 de la leçon précédente, il existe une base orthonormée $\mathbf{e}_1, \ldots, \mathbf{e}_n$ de $\mathcal{V}$ telle que

$$\mathbf{A}\mathbf{e}_q = \lambda_q \mathbf{e}_q \text{ pour tout } q = 1, \ldots, n.$$

Soit $\mathcal{P}_q = [\mathbf{e}_1, \ldots, \mathbf{e}_q]$ et soit $\mathcal{P}$ un sous-espace quelconque de dimension $n - q + 1$. Comme

$$\dim \mathcal{P}_q + \dim \mathcal{P} = q + (n - q + 1) = n + 1 > n,$$

le théorème 1 de la 1ère leçon entraîne $\mathcal{P}_q \cap \mathcal{P} \neq 0$, i.e. on trouve un vecteur $\mathbf{x} \neq 0$ de $\mathcal{P}_q \cap \mathcal{P}$. On estime, sans restreindre la généralité, que $(\mathbf{x}, \mathbf{x}) = 1$. Etant donné que $\mathbf{x} \in \mathcal{P}$, on a $\alpha(\mathcal{P}) \geqslant (\mathbf{A}\mathbf{x}, \mathbf{x})$. D'autre part, $\mathbf{x} \in \mathcal{P}_q$, donc $\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_q\mathbf{e}_q$, si bien que

$$\begin{aligned}(\mathbf{A}\mathbf{x}, \mathbf{x}) &= (\lambda_1 x_1 \mathbf{e}_1 + \ldots + \lambda_q x_q \mathbf{e}_q,\ x_1\mathbf{e}_1 + \ldots + x_q\mathbf{e}_q) = \\ &= \lambda_1 |x_1|^2 + \ldots + \lambda_q |x_q|^2 \geqslant \lambda_q (|x_1|^2 + \ldots + |x_q|^2) = \\ &= \lambda_q (\mathbf{x}, \mathbf{x}) = \lambda_q.\end{aligned}$$

Ainsi, $\alpha(\mathcal{P}) \geqslant \lambda_q$ pour tout sous-espace $\mathcal{P}$ de dimension $n - q + 1$ et, partant,

$$\inf \{\alpha(\mathcal{P});\ \dim \mathcal{P} = n - q + 1\} \geqslant \lambda_q.$$

Mais on a pour tout vecteur $\mathbf{x} = x_q\mathbf{e}_q + \ldots + x_n\mathbf{e}_n$ de $\mathcal{P}_{(q)} = [\mathbf{e}_q, \ldots, \mathbf{e}_n]$ de dimension $n - q + 1$

$$\begin{aligned}(\mathbf{A}\mathbf{x}, \mathbf{x}) &= \lambda_q |x_q|^2 + \ldots + \lambda_n |x_n|^2 \leqslant \\ &\leqslant \lambda_q (|x_q|^2 + \ldots + |x_n|^2) = \lambda_q (\mathbf{x}, \mathbf{x}) = \lambda_q,\end{aligned}$$

donc

$$\alpha(\mathcal{P}_{(q)}) \leqslant \lambda_q,$$

et

$$\inf \{\alpha(\mathcal{P});\ \dim \mathcal{P} = n - q + 1\} \leqslant \lambda_q. \ \square$$

Le résultat établi porte le nom de *propriété de maximum-minimum* des valeurs propres des opérateurs auto-adjoints.

L'idée de la démonstration du théorème 3 est évidente. Quant à la mener en détail, nous en laissons le soin au lecteur.

* * *

Dans un espace euclidien, chaque opérateur orthogonalement diagonalisable étant représenté dans une base orthonormée par une matrice diagonale (donc symétrique) est symétrique (auto-adjoint)

Ce fait démontre le

Théorème 4. *Un opérateur linéaire d'un espace euclidien est orthogonalement diagonalisable si et seulement s'il est symétrique.* □

Quant aux opérateurs auto-adjoints (hermitiens) d'un espace unitaire, ils ne constituent qu'une partie des opérateurs orthogonalement diagonalisables parce qu'une matrice hermitienne a nécessairement tous ses éléments diagonaux réels. Aussi un opérateur dont la matrice dans une base orthonormée est diagonale et possède au moins un élément non réel est orthogonalement diagonalisable sans être hermitien.

Définition 1. Un opérateur A d'un espace unitaire (ou euclidien) est dit *normal* s'il commute avec son adjoint.

On rappelle (voir leçon 19) que tout opérateur A d'un espace hermitien est représentable de façon unique sous la forme

$$A = B + iC,$$

B et C étant des opérateurs hermitiens.

Proposition 2. *Un opérateur* $A = B + iC$ *d'un espace hermitien est normal si et seulement si les opérateurs* B *et* C *commutent* ($BC = CB$).

Démonstration. Comme

$$A^* = B^* + (iC^*) = B^* - iC^* = B - iC,$$

on a

$$AA^* = (B + iC)(B - iC) = B^2 + C^2 + i(CB - BC)$$

et

$$A^*A = (B - iC)(B + iC) = B^2 + C^2 - i(CB - BC).$$

Par conséquent, une condition suffisante et nécessaire pour que $AA^* = A^*A$ est $CB - BC = 0$. □

On note qu'étant donné un opérateur A normal, l'opérateur $AA^* = A^*A$ est défini par la formule

$$AA^* = B^2 + C^2$$

qui est analogue à la formule donnant le carré du module d'un nombre complexe.

Si A est représenté dans une base orthonormée par A diagonale, la matrice dans la même base de son adjoint est la transposée de la conjuguée de A, donc diagonale. Comme deux matrices diagonales commutent nécessairement, il en est de même des opérateurs A et A^*, ce qui démontre que *tout opérateur orthogonalement diagonalisable d'un espace hermitien est normal.* □

Notre but immédiat est de prouver la réciproque. Voyons tout d'abord si l'on peut transposer les propositions 3 à 5 de la leçon 19 au cas d'opérateurs normaux.

Il est certes impossible de généraliser directement la proposition 3 vu que les valeurs propres (= racines caractéristiques) d'un opé-

rateur normal peuvent être des nombres complexes quelconques. Il y a plus. En cas hermitien, elle découle immédiatement de son analogue pour les opérateurs normaux qui s'énonce comme suit:

Proposition 3. *Etant donné un opérateur normal* **A**, *tout vecteur propre de* **A** *associé à une valeur propre* λ *est vecteur propre de* $\mathbf{A}^*$ *associé à la valeur propre* $\bar{\lambda}$.

Démonstration. Si **A** est normal, on a pour tout vecteur **x**:

$$(\mathbf{Ax}, \mathbf{Ax}) = (\mathbf{A}^*\mathbf{Ax}, \mathbf{x}) = (\mathbf{AA}^*\mathbf{x}, \mathbf{x}) = (\mathbf{A}^*\mathbf{x}, \mathbf{A}^*\mathbf{x}),$$

i.e.

$$|\mathbf{Ax}| = |\mathbf{A}^*\mathbf{x}|.$$

Chaque fois que l'opérateur **A** est normal, $\mathbf{A} - \lambda\mathbf{E}$ l'est également, si bien que la dernière égalité entraîne pour tout λ (du moment que $(\mathbf{A} - \lambda\mathbf{E})^* = \mathbf{A}^* - \bar{\lambda}\mathbf{E}$)

$$|(\mathbf{A} - \lambda\mathbf{E})\mathbf{x}| = |(\mathbf{A}^* - \bar{\lambda}\mathbf{E})\mathbf{x}|.$$

Aussi $(\mathbf{A}^* - \bar{\lambda}\mathbf{E})\mathbf{x} = 0$ dès que $(\mathbf{A} - \lambda\mathbf{E})\mathbf{x} = 0$. □

La proposition 4 de la leçon 19 reste en vigueur pour les opérateurs normaux sans qu'on y change quoi que ce soit, à savoir on a la

Proposition 4. *Etant donné un opérateur normal* **A**, *ses vecteurs propres quelconques* **x** *et* **y** *associés aux valeurs propres* λ, μ *distinctes sont orthogonaux.*

Démonstration. Si $\mathbf{Ax} = \lambda\mathbf{x}$, alors $(\mathbf{Ax}, \mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y})$. De même, si $\mathbf{Ay} = \mu\mathbf{y}$, donc $\mathbf{A}^*\mathbf{y} = \bar{\mu}\mathbf{y}$ par la proposition 3, alors $(\mathbf{x}, \mathbf{A}^*\mathbf{y}) = (\mathbf{x}, \bar{\mu}\mathbf{y}) = \mu(\mathbf{x}, \mathbf{y})$. Par conséquent, $\lambda(\mathbf{x}, \mathbf{y}) = (\mathbf{Ax}, \mathbf{y}) = (\mathbf{x}, \mathbf{A}^*\mathbf{y}) = \mu(\mathbf{x}, \mathbf{y})$ et, partant, $(\mathbf{x}, \mathbf{y}) = 0$ (car $\lambda \neq \mu$ par hypothèse). □

Quant à la proposition 5 de la leçon précédente, elle n'est en général pas vraie pour les opérateurs normaux. En effet, il existe des opérateurs normaux tels que le supplémentaire orthogonal des sous-espaces invariants associés soit non invariant (donner un exemple !). Cette proposition joue par contre pour les sous-espaces propres.

Proposition 5. *Le supplémentaire orthogonal* $\mathcal{P}_\lambda^\perp$ *d'un sous-espace propre quelconque* $\mathcal{P}_\lambda$ *d'un opérateur normal* **A** *est invariant par* **A**.

Démonstration. Si $\mathbf{x} \in \mathcal{P}_\lambda^\perp$, alors $(\mathbf{x}, \mathbf{y}) = 0$ pour tout $\mathbf{y} \in \mathcal{P}_\lambda$. Aussi $(\mathbf{Ax}, \mathbf{y}) = (\mathbf{x}, \mathbf{A}^*\mathbf{y}) = (\mathbf{x}, \bar{\lambda}\mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y}) = 0$ car $\mathbf{A}^*\mathbf{y} = \bar{\lambda}\mathbf{y}$ par la proposition 3. Donc $\mathbf{Ax} \in \mathcal{P}_\lambda^\perp$. □

On a déjà dit que cette propriété de **A** a été seule nécessaire pour démontrer la proposition 6 de la leçon précédente. Aussi la proposition persiste pour tout opérateur normal, ce qui garantit (prop. 4) l'existence d'une base orthonormée dans laquelle la matrice de l'opérateur est diagonale.

Ainsi, on vient de démontrer le

Théorème 5. *Un opérateur linéaire dans un espace hermitien est orthogonalement diagonalisable si et seulement s'il est normal.* □

Le théorème permet de ramener les propriétés d'un opérateur normal à celles de son spectre. Il est évident par exemple qu'*un opérateur normal* **A** *d'un espace hermitien est*

a) *hermitien*,

b) *inversible*,

c) *idempotent* (*i.e.* $\mathbf{A}^2 = \mathbf{A}$)

si et seulement si ses valeurs propres sont

a') *réelles*,

b') *non nulles*,

c') *égales à* 0 *ou à* 1

respectivement.

On note que les relations a)⇒a'), b)⇒b') et c)⇒c') ont lieu pour tout opérateur linéaire, tandis qu'elles ne sont vérifiées dans l'autre sens (cas le plus intéressant) que par les opérateurs normaux (donner des exemples !).

Il va de soi qu'on établit des résultats analogues pour les opérateurs symétriques dans les espaces euclidiens.

LEÇON 21

Opérateurs positifs. — Opérateurs isométriques. — Matrices unitaires. — Décomposition polaire des opérateurs inversibles. — Interprétation géométrique de la décomposition polaire. — Translations et transformations affines centrées. — Diagonalisation d'un opérateur unitaire. — Décomposition de la rotation de l'espace euclidien de dimension n en produit de rotations de plans bidimensionnels.

Proposition 1. *Les propriétés suivantes d'un opérateur linéaire* **A** *dans un espace* $\mathcal{T}$ *euclidien ou hermitien sont équivalentes*:

a) *Il existe un opérateur auto-adjoint* **B** *tel que*

$$\mathbf{A} = \mathbf{B}^2.$$

b) *Il existe un opérateur linéaire* **C** *tel que*

$$\mathbf{A} = \mathbf{C}^*\mathbf{C}.$$

c) *L'opérateur* **A** *est auto-adjoint et* $(\mathbf{Ax}, \mathbf{x}) \geqslant 0$ *quel que soit* $\mathbf{x} \in \mathcal{T}$.

d) *L'opérateur* **A** *est auto-adjoint et a toutes ses valeurs propres non négatives.*

L'équivalence est conservée si l'on renforce les propriétés en exigeant que **B** *et* **C** *de* a) *et* b) *soient inversibles, que* $(\mathbf{Ax}, \mathbf{x}) > 0$ *pour* $x \neq 0$ *dans* c) *et que toutes les valeurs propres soient positives* (*cas* d)).

Démonstration. a) $\Rightarrow$ b). Il suffit de poser $\mathbf{C} = \mathbf{B}$.

b) $\Rightarrow$ c). Si $\mathbf{A} = \mathbf{C}^*\mathbf{C}$, alors $(\mathbf{Ax}, \mathbf{x}) = (\mathbf{Cx}, \mathbf{Cx}) = |\mathbf{Cx}|^2 \geqslant 0$. Cela étant, l'inversibilité de **C** (i.e. $\mathbf{Cx} \neq 0$ pour $\mathbf{x} \neq 0$) entraîne $(\mathbf{Ax}, \mathbf{x}) > 0$ pour $\mathbf{x} \neq 0$.

c) $\Rightarrow$ d). Si $\mathbf{Ax} = \lambda\mathbf{x}$, alors $(\mathbf{Ax}, \mathbf{x}) = \lambda(\mathbf{x}, \mathbf{x})$, si bien que chaque fois que $(\mathbf{Ax}, \mathbf{x})$ est non négatif (resp. positif), λ est non négatif (resp. positif).

d) $\Rightarrow$ a). Soient $\mathbf{e}_1, \ldots, \mathbf{e}_n$ une base constituée de vecteurs propres de **A** et $\lambda_1, \ldots, \lambda_n$ les valeurs propres correspondantes. $\lambda_1 \geqslant 0, \ldots, \lambda_n \geqslant 0$ par hypothèse, si bien qu'on trouve (dans $\mathbb{R}$) les racines $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}$. On définit l'opérateur **B** par les égalités

$$\mathbf{Be}_1 = \sqrt{\lambda_1}\,\mathbf{e}_1, \ldots, \mathbf{Be}_n = \sqrt{\lambda_n}\,\mathbf{e}_n. \tag{1}$$

Il est clair que $\mathbf{B}^2 = \mathbf{A}$. □

Définition 1. Un opérateur **A** est dit *non négatif* s'il jouit des propriétés a) à d). Si **A** présente les propriétés a) à d) renforcées, il s'appelle un opérateur *positif*. Chaque opérateur auto-adjoint **B** vérifiant la relation $\mathbf{B}^2 = \mathbf{A}$ est par définition une *racine carrée*

de l'opérateur A. Une racine carrée non négative (ou positive) est désignée par $\sqrt{\mathbf{A}}$.

La formule (1) montre que *l'opérateur* $\sqrt{\mathbf{A}}$ *existe et est défini de façon unique pour tout opérateur* A *non négatif* (*ou positif*). □

Il est évident qu'*un opérateur non négatif est positif si et seulement s'il est inversible.* □

Une condition nécessaire et suffisante pour qu'un opérateur autoadjoint A d'un espace euclidien soit positif est la définie positivité de la fonctionnelle quadratique (Ax, x).

Nombreux sont les auteurs qui emploient une terminologie différente: un opérateur non négatif est pour eux un opérateur *positif*, un opérateur positif étant appelé *strictement positif*.

* * *

Les opérateurs positifs sont des analogues des nombres réels positifs. Voyons les opérateurs qui correspondent aux nombres complexes de module unité.

Proposition 2. *Les propriétés suivantes d'un opérateur linéaire* A *dans un espace* $\mathcal{V}$ *euclidien ou hermitien sont équivalentes*:

a) *On a pour tout* x *et tout* y *de* $\mathcal{V}$ *l'égalité*

$$(\mathbf{Ax}, \mathbf{Ay}) = (\mathbf{x}, \mathbf{y}).$$

b) *Tout vecteur* x *de* $\mathcal{V}$ *satisfait à l'égalité*

$$|\mathbf{Ax}| = |\mathbf{x}|.$$

c) *Etant donnée une base orthonormée* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *quelconque de* $\mathcal{V}$, *les vecteurs* $\mathbf{Ae}_1, \ldots, \mathbf{Ae}_n$ *constituent une base orthonormée de* $\mathcal{V}$.

d) *Les éléments* a_i^j *de la matrice représentant* A *dans une base orthonormée* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *quelconque de l'espace* $\mathcal{V}$ *vérifient les relations*

$$\sum_{k=1}^{n} a_i^k a_j^k = \delta_{ij}, \quad i, j = 1, \ldots, n \tag{2}$$

(*cas euclidien*) *et*

$$\sum_{k=1}^{n} a_i^k \overline{a_j^k} = \delta_{ij}, \quad i, j = 1, \ldots, n \tag{3}$$

(*cas hermitien*).

e) *On a l'égalité*

$$\mathbf{A}^*\mathbf{A} = \mathbf{E}.$$

f) *L'opérateur* A *est inversible et*

$$\mathbf{A}^{-1} = \mathbf{A}^*.$$

g) $\mathbf{AA^*} = \mathbf{E}$.

h) *Les éléments* a_i^j *de la matrice représentant* $\mathbf{A}$ *dans une base orthonormée* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *quelconque de* $\mathcal{T}$ *vérifient les relations*

$$\sum_{k=1}^{n} a_k^i a_k^j = \delta_{ij}, \quad i, j = 1, \ldots, n \tag{4}$$

(*cas euclidien*) *et*

$$\sum_{k=1}^{n} a_k^i \bar{a}_k^j = \delta_{ij}, \quad i, j = 1, \ldots, n \tag{5}$$

(*cas hermitien*).

Démonstration. On établira le diagramme suivant:

$$\begin{tikzcd}
a) \arrow[r, Leftrightarrow] \arrow[dr, Rightarrow] \arrow[d, Rightarrow] & e) \arrow[r, Leftrightarrow] & f) \arrow[r, Leftrightarrow] & g) \arrow[r, Leftrightarrow] & h) \\
 & b) \arrow[ur, Rightarrow] & & & \\
c) \arrow[r, Leftrightarrow] & d) \arrow[uur, Leftrightarrow] & & &
\end{tikzcd}$$

a) $\Rightarrow$ b). Il suffit de poser $\mathbf{y} = \mathbf{x}$.

a) $\Rightarrow$ c). Comme $(\mathbf{Ae}_i, \mathbf{Ae}_j) = (\mathbf{e}_i, \mathbf{e}_j)$, on a $(\mathbf{Ae}_i, \mathbf{Ae}_j) = \delta_{ij}$ pour $(\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij}$.

b) $\Rightarrow$ e). Si $|\mathbf{Ax}| = |\mathbf{x}|$, alors $((\mathbf{A^*A} - \mathbf{E})\,\mathbf{x}, \mathbf{x}) = (\mathbf{A^*Ax}, \mathbf{x}) - (\mathbf{x}, \mathbf{x}) = (\mathbf{Ax}, \mathbf{Ax}) - (\mathbf{x}, \mathbf{x}) = |\mathbf{Ax}|^2 - |\mathbf{x}|^2 = 0$, donc $\mathbf{A^*A} = \mathbf{E}$ (par suite de la symétrie de $\mathbf{A^*A} - \mathbf{E}$ pour le cas euclidien et par la proposition 1 de la leçon 19 dans le cas hermitien).

c) $\Leftrightarrow$ d). Par définition, $\mathbf{Ae}_i = a_i^j \mathbf{e}_j$, si bien que

$$(\mathbf{Ae}_i, \mathbf{Ae}_j) = \sum_{k=1}^{n} a_i^k a_j^k$$

(pour $\mathcal{T}$ euclidien) et

$$(\mathbf{Ae}_i, \mathbf{Ae}_j) = \sum_{k=1}^{n} a_i^k \bar{a}_j^k$$

(pour $\mathcal{T}$ hermitien). Par conséquent, c) $\Rightarrow$ d) et d) $\Rightarrow$ c).

a) $\Leftrightarrow$ e). Par définition, $(\mathbf{A^*Ax}, \mathbf{y}) = (\mathbf{Ax}, \mathbf{Ay})$. Aussi a) $\Rightarrow$ e) et e) $\Rightarrow$ a) (car l'égalité $(\mathbf{Cx}, \mathbf{y}) = (\mathbf{x}, \mathbf{y})$, avec un opérateur $\mathbf{C}$ et n'importe quels $\mathbf{x}$ et $\mathbf{y}$, a lieu si et seulement si $\mathbf{C} = \mathbf{E}$).

d) $\Leftrightarrow$ e) et g) $\Leftrightarrow$ h). L'opérateur $\mathbf{A^*}$ est représenté dans la base $\mathbf{e}_1, \ldots, \mathbf{e}_n$ par la matrice $(\bar{a}_i^j)$. La matrice de $\mathbf{AA^*}$ est donc formée de sommes $\sum_k a_k^i \bar{a}_k^j$ et la matrice de $\mathbf{A^*A}$ de sommes $\sum_k a_i^k \bar{a}_j^k$. Aussi d) $\Leftrightarrow$ e) et g) $\Leftrightarrow$ h).

e) $\Rightarrow$ f) et g) $\Rightarrow$ f). Voir 1° $\Rightarrow$ 5° et 3° $\Rightarrow$ 5° de la proposition 2 de la leçon 14.

f) $\Rightarrow$ e) et f) $\Rightarrow$ g). Ces relations ont lieu par définition. □

Définition 2. Un opérateur linéaire **A** sur un espace $\mathcal{T}$ euclidien ou hermitien est dit *isométrique* s'il jouit des propriétés a) à h). Les opérateurs isométriques s'appellent également opérateurs *orthogonaux* ou *unitaires* selon que $\mathcal{T}$ est euclidien ou hermitien.

La propriété a) signifie que **A** conserve les produits scalaires (et, en particulier, les angles), i.e. c'est *un homomorphisme* (voire un isomorphisme en vertu de f)) de $\mathcal{T}$ sur lui-même.

Chose à noter, *tout opérateur isométrique est normal* ($\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*$). □

* * *

Il est déjà apparu (I.14, prop. 4) que les matrices réelles ayant la propriété (2) ou (4) sont exactement les *matrices orthogonales*. Par analogie on appelle *matrices unitaires* les matrices à coefficients complexes présentant les propriétés (3) et (5). On a pour les matrices unitaires l'analogue suivant de la proposition 4 de I.14 (le symbole $\overline{A}^{\mathsf{T}}$ désigne une matrice transposée à éléments complexes conjugués).

Proposition 3. *Une matrice $A = (a_i^j)$ d'ordre n à coefficients complexes est unitaire si et seulement si elle possède l'une (donc chaque) propriété des propriétés équivalentes suivantes:*

a) *La matrice A est la matrice de passage d'une base orthonormée de l'espace hermitien de dimension n à une autre base orthonormée de cet espace.*

b) *Les colonnes de A forment une famille orthonormée de vecteurs de l'espace hermitien $\mathbb{C}^n$.*

c) $\overline{A}^{\mathsf{T}}A = E$.

d) *La matrice A est inversible et*

$$A^{-1} = \overline{A}^{\mathsf{T}}.$$

e) $A\,\overline{A}^{\mathsf{T}} = E$.

f) *Les lignes de A forment une famille orthonormée de vecteurs de l'espace hermitien $\mathbb{C}^n$.*

Démonstration. Soit un opérateur linéaire **A** de matrice A dans une base orthonormée. Les propriétés a) à f) deviennent les propriétés c) à h) de la proposition 2. □

Comme $\det \overline{A}^{\mathsf{T}} = \overline{\det A}$, les affirmations c) et e) impliquent

$$|\det A| = 1$$

quelle que soit A unitaire.

Les matrices unitaires d'ordre n forment manifestement un groupe appelé *groupe unitaire* et noté U (n). Le sous-groupe de U (n) constitué de matrices unimodulaires (i.e. de matrices de déterminant 1) est le groupe spécial unitaire SU (n).

* * *

Proposition 4. *Tout opérateur inversible* **A** *dans un espace euclidien (ou hermitien) se décompose de façon unique en produit d'un opérateur isométrique* **U** *et d'un opérateur positif* **P**:

$$\mathbf{A} = \mathbf{PU}. \tag{6}$$

Démonstration. Selon la proposition 1, l'opérateur $\mathbf{A}^*\mathbf{A}$ est positif, si bien qu'il existe une racine carrée positive

$$\mathbf{P} = \sqrt{\mathbf{A}^*\mathbf{A}}.$$

Soit $\mathbf{U} = \mathbf{AP}^{-1}$, auquel cas $\mathbf{U}^* = (\mathbf{P}^*)^{-1}\mathbf{A}^* = \mathbf{P}^{-1}\mathbf{A}^*$ (car **P** est auto-adjoint) et $\mathbf{U}^*\mathbf{U} = \mathbf{P}^{-1}\mathbf{A}^*\mathbf{AP}^{-1} = \mathbf{P}^{-1}\mathbf{P}^2\mathbf{P}^{-1} = \mathbf{E}$. Ainsi, $\mathbf{A} = \mathbf{UP}$, **U** étant isométrique et **P** positif.

Si $\mathbf{UP} = \mathbf{VQ}$, avec **U**, **V** des opérateurs isométriques et **P**, **Q** des opérateurs positifs, alors $\mathbf{PU}^* = \mathbf{QV}^*$, donc

$$\mathbf{P}^2 = \mathbf{PU}^*\mathbf{UP} = \mathbf{QV}^*\mathbf{VQ} = \mathbf{Q}^2.$$

Par conséquent, $\mathbf{P} = \mathbf{Q}$ (une racine carrée positive a une valeur unique) et, partant, $\mathbf{U} = \mathbf{V}$, ce qui prouve l'unicité de (6). □

La décomposition (6) est connue sous le nom de *décomposition polaire* de l'opérateur **A**. C'est un analogue de $re^{i\varphi} = r(\cos\varphi + i\sin\varphi)$, i.e. de la décomposition d'un nombre complexe arbitraire en produit de son module r et du nombre $e^{i\varphi}$ de module unité.

* * *

On rappelle (voir I.26) qu'une *transformation affine* d'un espace affine $\mathcal{A}$ est par définition un automorphisme arbitraire de $\mathcal{A}$, i.e. une transformation définie par égalité des coordonnées dans deux repères affines. Si l'on choisit dans $\mathcal{A}$ une origine O, un point de rayon vecteur **x** devient par toute transformation affine un point de rayon vecteur

$$\mathbf{y} = \mathbf{Ax} + \mathbf{b}, \tag{7}$$

A étant un opérateur linéaire inversible et **b** un vecteur fixe (en fait, ce n'est rien d'autre que la formule (2) de I.27).

De même, on appelle *transformation orthogonale* d'un espace ponctuel euclidien $\mathcal{E}$ une transformation de $\mathcal{E}$ définie par égalité des coordonnées dans deux systèmes de coordonnées (rectangulaires) euclidiennes. On a une fois de plus la formule (7) à la différence que **A** est un opérateur orthogonal.

On introduit de même les *espaces ponctuels hermitiens* $\mathscr{E}$ considérés comme espaces affines dont les espaces vectoriels attachés sont structurés en espaces vectoriels hermitiens. Leurs automorphismes sont des *transformations unitaires* mises sous forme (7), avec **A** un opérateur unitaire.

Tout espace ponctuel euclidien (ou hermitien) étant affine, il y a lieu de parler de ses transformations (7) affines. Il correspond alors à la décomposition polaire $\mathbf{A} = \mathbf{UP}$ la décomposition de (7) en produit de la transformation affine

$$\mathbf{y} = \mathbf{Px} \tag{8}$$

et de la transformation orthogonale (ou unitaire)

$$\mathbf{y} = \mathbf{Ux} + \mathbf{b}.$$

La transformation (8) s'écrit en coordonnées rectangulaires convenables :

$$\begin{aligned} y_1 &= \lambda_1 x_1, \\ &\cdots\cdots \\ y_n &= \lambda_n x_n, \end{aligned}$$

$\lambda_1 > 0, \ldots, \lambda_n > 0$. C'est donc le produit de n contractions relativement à n hyperplans perpendiculaires entre eux. Ainsi, on dit que *toute transformation affine de l'espace ponctuel euclidien* (resp. *hermitien*) *de dimension n est le produit d'une transformation orthogonale* (resp. *unitaire*) *et de n contractions relativement à n hyperplans perpendiculaires entre eux.* □

On retrouve pour $n = 2$ la proposition 1 de I.27.

* * *

Quand $\mathbf{A} = \mathbf{E}$, la transformation (7) s'écrit

$$\mathbf{y} = \mathbf{x} + \mathbf{b}$$

et constitue par définition une *translation* de vecteur **b**. Si $\mathbf{b} = \mathbf{0}$, la transformation (7) devient

$$\mathbf{y} = \mathbf{Ax}.$$

C'est une *transformation affine centrée*. Elle laisse fixe le point O appelé *centre* de la transformation. Toute transformation affine se ramène à une translation suivie d'une transformation affine centrée.

On souligne que (7) avec $\mathbf{b} \neq \mathbf{0}$ peut être une transformation affine centrée (de centre autre que le point O). Une condition nécessaire et suffisante en est l'existence d'un vecteur $\mathbf{x}_0$ (rayon vecteur du centre) tel que

$$\mathbf{x}_0 = \mathbf{Ax}_0 + \mathbf{b},$$

i.e. $(\mathbf{A} - \mathbf{E})\,\mathbf{x}_0 = \mathbf{b}$. En particulier, cela arrive nécessairement pour $\mathbf{A} - \mathbf{E}$ *inversible*, i.e. pour $\mathbf{A}$ n'ayant pas de valeur propre 1.

* * *

Une transformation orthogonale qui est une transformation affine centrée s'appelle *rotation généralisée*. C'est une *rotation* tout court si l'opérateur orthogonal $\mathbf{A}$ a son déterminant $= +1$ (i.e. si $\mathbf{A}$ conserve l'orientation).

Pour avoir une idée des rotations, on doit se familiariser avec les opérateurs orthogonaux. Il y a intérêt à faire précéder leur étude par celle des opérateurs unitaires.

Proposition 5. *Le spectre d'un opérateur unitaire quelconque* $\mathbf{A}$ *est dans le plan de la variable complexe sur la circonférence unité, i.e. chaque racine caractéristique* λ *de* $\mathbf{A}$ *est de module unité*:

$$|\lambda| = 1.$$

Démonstration. Toute racine caractéristique λ sur $\mathbb{C}$ est valeur propre, i.e. on trouve un vecteur $\mathbf{x}_0 \neq 0$ tel que $\mathbf{A}\mathbf{x}_0 = \lambda\mathbf{x}_0$, auquel cas

$$(\mathbf{x}_0, \mathbf{x}_0) = (\mathbf{A}\mathbf{x}_0, \mathbf{A}\mathbf{x}_0) = (\lambda\mathbf{x}_0, \lambda\mathbf{x}_0) = \lambda\bar{\lambda}\,(\mathbf{x}_0, \mathbf{x}_0),$$

donc $\lambda\bar{\lambda} = 1$. □

Théorème 1. *Etant donné un opérateur unitaire arbitraire* $\mathbf{A}$, *il existe une base orthonormée dans laquelle il est représenté par la matrice*

$$\begin{pmatrix} e^{i\varphi_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\varphi_n} \end{pmatrix}.$$

Démonstration. Un opérateur unitaire étant normal est orthogonalement diagonalisable. Cette propriété jointe à la proposition 5 fournit le résultat voulu. □

* * *

Soit $\mathbf{A}$ un opérateur orthogonal dans $\mathcal{V}$ euclidien (réel).

On introduit son complexifié

$$\mathbf{A}^{\mathbb{C}}(\mathbf{x} + i\mathbf{y}) = \mathbf{A}\mathbf{x} + i\mathbf{A}\mathbf{y}$$

qui est un opérateur linéaire sur le complexifié

$$\mathcal{V}^{\mathbb{C}} = \mathcal{V} + i\mathcal{V}$$

de l'espace $\mathcal{V}$ (voir leçon 17).

Etant donnés

$$\mathbf{z} = \mathbf{x} + i\mathbf{y} \in \mathcal{V}^{\mathbb{C}}, \quad \mathbf{z}_1 = \mathbf{x}_1 + i\mathbf{y}_1 \in \mathcal{V}^{\mathbb{C}}$$

quelconques, on pose

$$(z, z_1) = [(x, x_1) + (y, y_1)] - i\,[(x, y_1) - (x_1, y)].$$

On vérifie de façon banale que la fonctionnelle $z, z_1 \mapsto (z, z_1)$ est sesquilinéaire hermitienne définie positive, i.e. on l'accepte en qualité de produit scalaire défini sur le complexifié $\mathcal{V}^{\mathbb{C}}$. L'espace $\mathcal{V}^{\mathbb{C}}$ est donc un espace hermitien pour cette multiplication.

Comme

$$(A^{\mathbb{C}}z, A^{\mathbb{C}}z_1) =$$
$$= [(Ax, Ax_1) + (Ay, Ay_1)] - i\,[(Ax, Ay_1) - (Ax_1, Ay)] =$$
$$= [(x, x_1) + (y, y_1)] - i\,[(x, y_1) - (x_1, y)] = (z, z_1),$$

le complexifié $A^{\mathbb{C}}$ *de l'opérateur orthogonal* A *est unitaire.* Il en résulte en particulier la propriété de $A^{\mathbb{C}}$ d'être diagonalisable.

On en déduit (voir leçon 17, th. 1) qu'il existe une base de $\mathcal{V}$ dans laquelle la matrice de l'opérateur A est somme directe de matrices d'ordre 1 de la forme λ et de matrices d'ordre 2

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Cela étant, les nombres réels λ sont racines caractéristiques de $A^{\mathbb{C}}$, donc $|\lambda| = 1$, i.e. $\lambda = \pm 1$. Quant à α, β, c'est respectivement la partie réelle et le coefficient de la partie imaginaire de la racine caractéristique non réelle $\lambda = e^{i\varphi}$ du complexifié $A^{\mathbb{C}}$, si bien que $\alpha = \cos\varphi$ et $\beta = \sin\varphi$, $-\pi < \varphi \leqslant \pi$ et $\varphi \neq 0$.

Puisque les matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

s'écrivent

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \tag{9}$$

(pour $\varphi = 0$ et $\varphi = \pi$ respectivement), on dit que *la matrice associée à un opérateur orthogonal dans une base* $e_1, \ldots, e_n$ *de l'espace* $\mathcal{V}$ *est pour* $n = 2m + 1$ *somme directe de* m *matrices* (9), *avec* $-\pi < \varphi \leqslant \pi$, *et d'une matrice d'ordre* 1 (± 1) *et que si* $n = 2m$, *elle est, ou bien somme directe de* m *matrices* (9), *ou bien somme directe de* $m - 1$ *matrices* (9) *et d'une matrice*

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad \square$$

La base $e_1, \ldots, e_n$ de $\mathcal{V}$ s'obtient (voir leçon 17) d'une base $e_1^{\mathbb{C}}, \ldots, e_n^{\mathbb{C}}$ de $\mathcal{V}^{\mathbb{C}}$ qui jouit des propriétés suivantes:

a) chaque vecteur $e_q^{\mathbb{C}}$ est vecteur propre de l'opérateur unitaire $A^{\mathbb{C}}$;

b) si la valeur propre $\lambda_q = e^{i\varphi_q}$ associée à un vecteur propre $e_q^{\mathbb{C}}$ est réelle (i.e. si $\varphi_q = 0, \pi$), alors il en est de même de $e_q^{\mathbb{C}}$, et si $0 < \varphi_q < \pi$, alors le vecteur $e_{q+1}^{\mathbb{C}}$ est complexe conjugué de $e_q^{\mathbb{C}}$ correspondant à la valeur propre complexe conjuguée $\bar{\lambda}_q = e^{-i\varphi_q}$.

Cela étant, on a

$$e_q^{\mathbb{C}} = \begin{cases} e_q & \text{si } \varphi_q = 0, \pi, \\ e_q + ie_{q+1} & \text{si } 0 < \varphi_q < \pi, \\ e_{q-1} - ie_q & \text{si } -\pi < \varphi_q < 0. \end{cases}$$

On fait l'hypothèse complémentaire de $e_1^{\mathbb{C}}, \ldots, e_n^{\mathbb{C}}$ orthonormée ($A^{\mathbb{C}}$ étant orthogonalement diagonalisable). Comme

$$e_q = \begin{cases} e_q^{\mathbb{C}} & \text{pour } \varphi_q = 0, \pi, \\ \dfrac{e_q^{\mathbb{C}} + e_{q+1}^{\mathbb{C}}}{2} & \text{pour } 0 < \varphi < \pi, \\ \dfrac{e_{q-1}^{\mathbb{C}} - e_q^{\mathbb{C}}}{2i} & \text{pour } -\pi < \varphi_q < 0, \end{cases}$$

on a les égalités

$$(e_p, e_q) = 0 \text{ si } p \neq q,$$

$$(e_p, e_p) = \begin{cases} 1 & \text{si } \varphi_p = 0, \pi, \\ 2 & \text{si } \varphi_p \neq 0, \pi. \end{cases}$$

Si l'on multiplie tous les vecteurs e_p, $\varphi_p \neq 0, \pi$, par $\sqrt{2}$, on obtient donc une base orthonormée. La matrice de A étant conservée par cette opération (on le vérifie aisément), on vient de démontrer le

Théorème 2. *Pour tout opérateur orthogonal* A *de l'espace euclidien* $\mathcal{V}$ *de dimension* n, *il existe une base orthonormée dans laquelle la matrice de* A *s'écrit pour* $n = 2m + 1$

$$(10) \quad \begin{pmatrix} \varepsilon & & & & \\ & \begin{matrix} \cos\varphi_1 & \sin\varphi_1 \\ -\sin\varphi_1 & \cos\varphi_1 \end{matrix} & & 0 & \\ & & \begin{matrix} \cos\varphi_2 & \sin\varphi_2 \\ -\sin\varphi_2 & \cos\varphi_2 \end{matrix} & & \\ & 0 & & \ddots & \\ & & & & \begin{matrix} \cos\varphi_m & \sin\varphi_m \\ -\sin\varphi_m & \cos\varphi_m \end{matrix} \end{pmatrix},$$

$\varepsilon = \pm 1$, *et pour* $n = 2m$, *ou bien*

$$(11)\quad \begin{pmatrix} \begin{vmatrix} \cos\varphi_1 & \sin\varphi_1 \\ -\sin\varphi_1 & \cos\varphi_1 \end{vmatrix} & & & 0 \\ & \begin{vmatrix} \cos\varphi_2 & \sin\varphi_2 \\ -\sin\varphi_2 & \cos\varphi_2 \end{vmatrix} & & \\ & & \ddots & \\ 0 & & & \begin{vmatrix} \cos\varphi_m & \sin\varphi_m \\ -\sin\varphi_m & \cos\varphi_m \end{vmatrix} \end{pmatrix},$$

ou bien

$$(12)\quad \begin{pmatrix} \begin{vmatrix} \cos\varphi_1 & \sin\varphi_1 \\ -\sin\varphi_1 & \cos\varphi_1 \end{vmatrix} & & & 0 \\ & \ddots & & \\ & & \begin{vmatrix} \cos\varphi_{m-1} & \sin\varphi_{m-1} \\ -\sin\varphi_{m-1} & \cos\varphi_{m-1} \end{vmatrix} & \\ 0 & & & \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} \end{pmatrix}. \quad \square$$

On note que le déterminant de (10), (11) et (12) est égal à ε, à 1 et à -1 respectivement.

Le théorème 2 énoncé en termes de transformations orthogonales d'espaces ponctuels dit que *toute rotation de l'espace euclidien de dimension n est la composée de rotations dans $m = \left[\frac{n}{2}\right]$ plans bidimensionnels perpendiculaires entre eux et que toute rotation généralisée changeant l'orientation se ramène à une rotation possédant un axe (i.e. une droite dont tous les points restent fixes) suivie d'une symétrie dans un hyperplan perpendiculaire à l'axe de rotation.* Toute rotation a un axe pour $n = 2m + 1$, tandis que ce n'est pas toujours le cas pour $n = 2m$ (on cite comme exemple les rotations (11), où $\varphi_p \neq \neq 0, \pi$ quel que soit $p = 1, \ldots, m$).

Comme une rotation sans axe (ou, plus précisément, l'opérateur orthogonal associé de l'espace vectoriel attaché) n'a pas de valeurs propres égales à 1, sa composée avec toute translation est encore une rotation, mais le centre de rotation n'est plus le même. On a une affirmation analogue pour les rotations d'axe si le vecteur translation n'est parallèle à aucun axe (car ils peuvent être plusieurs). Il en résulte que *tout déplacement de l'espace euclidien est un déplacement hélicoïdal*, i.e. le produit d'une rotation et d'une translation de vecteur parallèle à un axe de rotation.

LEÇON 22

Fonctions différentiables. — Hypersurfaces différentiables. — Gradient. — Dérivées par rapport à un vecteur. — Champs vectoriels. — Points singuliers des champs vectoriels. — Champs vectoriels en tant que modules. — Champs de potentiel et champs irrotationnels. — Rotationnel. — Divergence des champs vectoriels. — Analyse vectorielle. — Hamiltonien. — Formules de produits. — Opérateurs composés.

En plus que l'espace $\mathbb{R}^n$ des vecteurs lignes est un modèle numérique d'espaces affines ou euclidiens de dimension n, il est le domaine de définition des fonctions $F(x_1, \ldots, x_n)$ de n variables. Ici la Géométrie et l'Analyse mathématique (la théorie des fonctions) chevauchent au point de faire corps. On ne s'étonne donc pas si c'est à l'Analyse qu'on doit l'une des premières définitions (ou *explications*) rigoureuses (et les plus importantes) des notions intuitives d'une courbe dans le plan, d'une surface dans l'espace de dimension 3 et, en général, d'une hypersurface dans l'espace n-dimensionnel.

Cette définition s'appuie sur l'idée d'une hypersurface (ou d'une courbe pour $n = 2$) comme du « lieu géométrique » des points dont les coordonnées satisfont à une condition de la forme

$$F(x_1, \ldots, x_n) = 0. \tag{1}$$

Comme il s'agit d'une courbe ou d'une surface « lisse » sans points anguleux, on suppose naturellement la fonction F *différentiable de classe* C^∞, i.e. F admet des dérivées partielles (forcément continues) de tous ordres. Dans la pratique (y compris dans la démonstration des théorèmes), on utilise essentiellement les dérivées de deux premiers ordres, si bien que F sera partout supposée admettre des dérivées partielles continues jusqu'à un ordre $k \geqslant 1$ (nous resterons donc fidèles à un principe mathématique général selon lequel on doit faire autant d'hypothèses qu'on s'en propose d'utiliser). D'autre part, l'obligation de se surveiller constamment afin d'éviter tout recours aux dérivées d'ordre supérieur pousse à ne pas préciser k, i.e. à exiger que toutes les fonctions possèdent des dérivées partielles continues de tout ordre voulu. Nous dirons de ces fonctions qu'elles sont *différentiables*.

La condition de fonction différentiable étant locale peut ne pas être remplie en certains points. Aussi nous considérerons les équations (1) non dans $\mathbb{R}^n$ tout entier, mais dans un ensemble $U \subset \mathbb{R}^n$ ouvert (disons, dans une boule ouverte). L'ensemble des fonctions

$\mathbf{x} \longmapsto F(\mathbf{x})$ définies et différentiables en tout point $\mathbf{x} = (x_1, \ldots, x_n) \in U$ sera désigné par $\mathcal{F}(U)$. Il est évidemment un anneau et un espace vectoriel (de dimension infinie) sur le corps $\mathbb{R}$.

* * *

S'agissant des fonctions différentiables simples (tels les polynômes), les ensembles définis par la condition (1) répondent de règle à l'idée intuitive de ce que c'est qu'une surface, bien qu'on se limite des fois à un ouvert de $\mathbb{R}^n$. Aussi on a longtemps tenu les ensembles définis par (1), avec F une fonction différentiable, comme plus ou moins aptes à jouer le rôle d'hypersurfaces (de courbes si $n = 2$). Quelle surprise ce fut pour tous quand le mathématicien américain Whitney démontra il y a une quarantaine d'années le théorème suivant : *étant donné un ensemble fermé $C \subset \mathbb{R}^n$ quelconque, on trouve sur $\mathbb{R}^n$ une fonction F différentiable (de classe C^∞) telle que $F(\mathbf{x}) = 0$ si et seulement si $\mathbf{x} \in C$.* (On constate sans peine que la condition de C fermé est nécessaire. Mais voilà qu'elle est suffisante !) Ce théorème sera démontré au cours du troisième semestre, et nous nous bornerons pour l'instant à un exemple.

Exemple. Une fonction F définie par les égalités :

$$F(\mathbf{x}) = \begin{cases} 0 & \text{si } |\mathbf{x}| \leqslant 1 ; \\ e^{\frac{1}{|\mathbf{x}|-1}} & \text{si } |\mathbf{x}| > 1, \end{cases}$$

avec $|\mathbf{x}| = \sqrt{x_1^2 + \ldots + x_n^2}$, est de classe C^∞ sur $\mathbb{R}^n$ tout entier. D'autre part, l'ensemble des points $\mathbf{x} \in \mathbb{R}^n$ tels que $F(\mathbf{x}) = 0$ forme la boule (le disque pour $n = 2$) $|\mathbf{x}| \leqslant 1$.

Avec le théorème de Whitney, on comprend la raison des restrictions complémentaires ajoutées à la condition de F différentiable. On apprend en Analyse que la *condition de régularité* est que le vecteur

$$\operatorname{grad} F = \left(\frac{\partial F}{\partial x_1}, \ldots, \frac{\partial F}{\partial x_n}\right)$$

(le *gradient* de la fonction F) soit non nul en tout point de l'hypersurface (1), i.e. qu'au moins l'une des dérivées partielles

$$\frac{\partial F}{\partial x_1}, \ldots, \frac{\partial F}{\partial x_n} \tag{2}$$

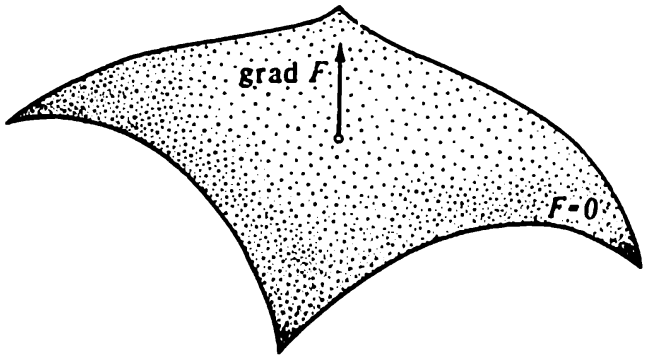
diffère de 0. Cela nous conduit à la

Définition 1. L'ensemble $\mathcal{H}$ des points $\mathbf{x} = (x_1, \ldots, x_n)$ d'un ouvert $U \subset \mathbb{R}^n$ tels qu'on ait

$$F(\mathbf{x}) = 0, \tag{3}$$

F étant une fonction différentiable sur U, s'appelle *hypersurface différentiable* (ou *régulière*) dans U si au moins une dérivée partielle

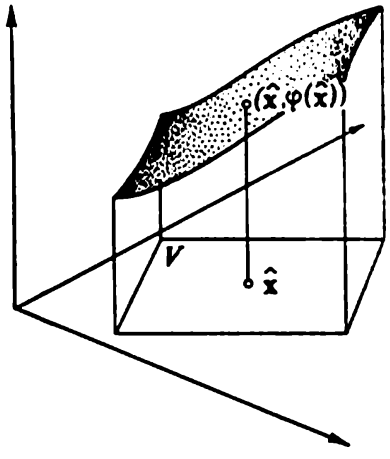
(2) est non nulle en chaque point $\mathbf{x} \in \mathcal{H}$.



Les points de $\mathbb{R}^{n-1}$ seront désignés par $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, . . ., le symbole $\hat{\mathbf{x}}$ représentant pour tout $\mathbf{x} = (x_1, \ldots, x_{n-1}, x_n)$ de $\mathbb{R}^n$ le point $(x_1, \ldots, x_{n-1})$ de $\mathbb{R}^{n-1}$. Etant donné un ensemble quelconque C de $\mathbb{R}^n$, on notera de même $\hat{C}$ l'ensemble de tous les points $\hat{\mathbf{x}} \in \mathbb{R}^{n-1}$, le point $\mathbf{x}$ appartenant à C. On écrira $\mathbf{x} = (\hat{\mathbf{x}}, x_n)$ au lieu de $\mathbf{x} = (x_1, \ldots, x_{n-1}, x_n)$.

On rappelle que le *graphe* d'une fonction différentiable $x_n = \varphi(\hat{\mathbf{x}})$ définie sur un ouvert $V \subset \mathbb{R}^{n-1}$ est par définition l'ensemble de tous les points $(\hat{\mathbf{x}}, \varphi(\hat{\mathbf{x}})) \in \mathbb{R}^n$. *Tout graphe est* évidemment *une hypersurface différentiable* telle que $U = V \times \mathbb{R} \subset \mathbb{R}^n$ et $F(\mathbf{x}) = \varphi(\hat{\mathbf{x}}) - x_n$ car $\frac{\partial F}{\partial x_n}(\mathbf{x}) = -1$ pour tout $\mathbf{x} \in U$. □

La réciproque est manifestement fausse. Par exemple, la circonférence $x^2 + y^2 = 1$ du plan n'est jamais le graphe d'une fonction, tandis qu'elle l'est au voisinage de chacun de ses points (ainsi, elle représente $y = \sqrt{1 - x^2}$ au voisinage de $(0, 1)$, la fonction $y = -\sqrt{1 - x^2}$ au voisinage de $(0, -1)$ et $x = \sqrt{1 - y^2}$ autour de $(1, 0)$, auquel cas le rôle de x_n incombe non à y, mais à x).



Graphe d'une fonction différentiable

Il se trouve qu'une affirmation analogue est vraie pour toute hypersurface $\mathcal{H}$, i.e. on dit qu'*au voisinage de chacun de ses points, toute hypersurface* (3) *est* à un changement de coordonnées près *le graphe d'une fonction différentiable.* Cette affirmation constitue le contenu géométrique d'un résultat connu en Analyse sous le nom de

Théorème des fonctions implicites. *Soient* $U \subset \mathbb{R}^n$ *un ouvert,* $\mathbf{x}_0 = (x_1^{(0)}, \ldots, x_n^{(0)})$ *un point de* U *et* $F: U \to \mathbb{R}$ *une fonction différentiable sur* U (i.e. *dans* $\mathcal{F}(U)$) *telle que*

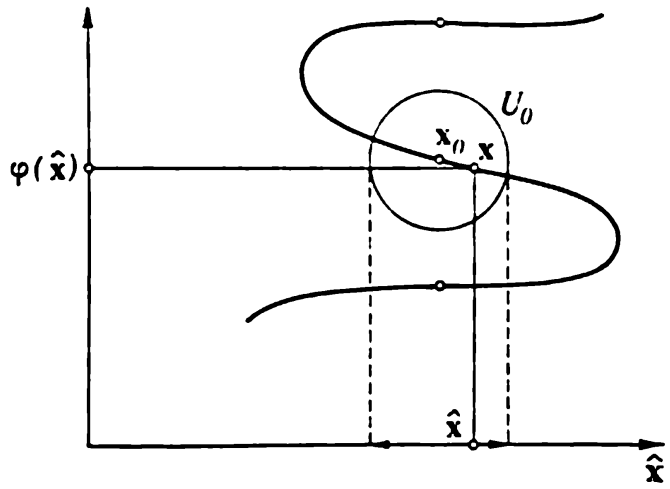
$$F(\mathbf{x}_0) = 0 \quad et \quad \frac{\partial F}{\partial x_n}(\mathbf{x}_0) \neq 0.$$

Il existe dans $\mathbb{R}^n$ *un voisinage* $U_0 \subset U$ *de* $\mathbf{x}_0$ *et une fonction* $x_n = \varphi(\hat{\mathbf{x}})$ *définie et différentiable dans le voisinage* $\hat{U}_0 \subset \mathbb{R}^{n-1}$ *de* $\hat{\mathbf{x}}_0 = (x_1^{(0)}, \ldots, x_{n-1}^{(0)})$ *qui jouissent des propriétés suivantes:*

a) $\varphi(\hat{\mathbf{x}}_0) = x_n^{(0)}$;

b) $(\hat{\mathbf{x}}, \varphi(\hat{\mathbf{x}})) \in U_0$ *pour tout point* $\hat{\mathbf{x}} \in \hat{U}_0$;

c) *si* $\mathbf{x} = (\hat{\mathbf{x}}, x_n) \in U_0$, *alors* $F(\mathbf{x}) \neq 0$ *sous la condition nécessaire et suffisante* $x_n = \varphi(\hat{\mathbf{x}})$. □



Il paraît que les graphes des fonctions différentiables d'une et de deux variables répondent à l'idée intuitive qu'on se fait des courbes et des surfaces « lisses », si bien que, d'après le théorème énoncé, la définition 1 ne contredit en tout cas pas l'intuition. D'autre part, la classe des hypersurfaces différentiables est suffisamment vaste pour que le fait de la former soit justifié.

* * *

Le fait de se borner à $\mathbb{R}^n$ est sans importance. En effet, la notion d'hypersurface différentiable est généralisée par un isomorphisme de coordonnées $\mathcal{A}^n \to \mathbb{R}^n$ à un espace affine (ou euclidien) quel-

conque $\mathcal{A}^n$ de dimension n, et elle reste intrinsèque (i.e. indépendante de l'isomorphisme de coordonnées).

Quant à la notion de gradient, elle reste intrinsèque (dans $\mathcal{A}^n$) si (et seulement si) dans tout changement de coordonnées

$$(4)\qquad \begin{aligned} x_1 &= c_{11}y_1 + \ldots + c_{n1}y_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= c_{1n}y_1 + \ldots + c_{nn}y_n \end{aligned}$$

les dérivées partielles (2) se transforment comme composantes d'un vecteur. Or, on constate aisément qu'il n'en est pas ainsi.

En effet, la fonction $F(x_1, \ldots, x_n)$ devient par (4) la fonction

$$G(y_1, \ldots, y_n) = \\ = F(c_{11}y_1 + \ldots + c_{n1}y_n, \ldots, c_{1n}y_1 + \ldots + c_{nn}y_n),$$

et la règle de dérivation d'une fonction composée entraîne

$$\frac{\partial G}{\partial y_i} = \sum_{j=1}^{n} c_{ij} \frac{\partial F}{\partial x_j},$$

i.e. dans le changement de coordonnées (4), les dérivées partielles (2) se transforment comme coordonnées d'un covecteur (voir leçon 4). Ainsi, *on assimile* grad F *à un covecteur.*

Tout cours d'Analyse suppose tacitement $\mathbb{R}^n$ être un espace euclidien muni du produit scalaire usuel $(\mathbf{x}, \mathbf{y}) = x_1y_1 + \ldots + x_ny_n$, si bien qu'on y identifie covecteurs et vecteurs. On ne saurait toutefois oublier que le gradient est « bel et bien » un covecteur sinon on risque de commettre des erreurs (et on les commet effectivement).

* * *

Les dérivées partielles sont un cas spécial de *dérivées par rapport à un vecteur* qu'on définit pour tout vecteur $\mathbf{k}$ de $\mathbb{R}^n$ par la formule

$$\frac{\partial F}{\partial \mathbf{k}}(\mathbf{x}) = \lim_{t \to 0} \frac{F(\mathbf{x} + t\mathbf{k}) - F(\mathbf{x})}{t},$$

i.e.

$$\frac{\partial F}{\partial \mathbf{k}}(\mathbf{x}) = \left.\frac{dF(\mathbf{x} + t\mathbf{k})}{dt}\right|_{t=0}.$$

Si $\mathbf{k} = (k_1, \ldots, k_n)$, on a par suite de la règle de dérivation d'une fonction composée :

$$\frac{\partial F}{\partial \mathbf{k}} = k_1 \frac{\partial F}{\partial x_1} + \ldots + k_n \frac{\partial F}{\partial x_n},$$

i.e.

$$\frac{\partial F}{\partial \mathbf{k}} = (\mathbf{k}, \text{ grad } F). \tag{5}$$

D'ailleurs, on ne s'intéresse d'ordinaire qu'au cas de $|\mathbf{k}| = 1$, i.e. à $\mathbf{k}$ un vecteur unitaire. On dit alors de $\frac{\partial F}{d\mathbf{k}}$ que c'est une dérivée de F *dans la direction du vecteur* $\mathbf{k}$, auquel cas les dérivées partielles ne sont autres que les dérivées dans la direction des axes de coordonnées.

Selon (5), le nombre $\frac{\partial F}{\partial \mathbf{k}}$ atteint son maximum (pour $|\mathbf{k}| = 1$) quand $\mathbf{k}$ est le vecteur unitaire de grad F. On dit donc que *le vecteur* grad F *a la direction de plus grande croissance de* F.

On note que le produit scalaire de la formule (5) ne signifie nullement qu'on est en cas euclidien. En effet, le second membre de (5) n'est évidemment rien d'autre que la valeur sur le vecteur $\mathbf{k}$ de grad F considéré comme covecteur. Quant à la dérivée $\frac{\partial F}{\partial \mathbf{k}}$, on la définit en général sans supposer que l'espace est euclidien.

* * *

On conçoit que grad F varie en général d'un point à l'autre, i.e. c'est une fonction vectorielle sur U. On donne à ces fonctions le nom de champs vectoriels. Voici la définition générale.

Définition 2. Chaque famille X de n fonctions

$$\mathbf{x} \longmapsto X_i(\mathbf{x}), \quad i = 1, \ldots, n,$$

avec $\mathbf{x} = (x_1, \ldots, x_n) \in U$, s'appelle un *champ vectoriel* sur U. Un champ vectoriel est *différentiable* s'il en est ainsi de toutes les fonctions X_i.

Du point de vue formel, un champ vectoriel sur U n'est rien d'autre qu'une application différentiable $U \to \mathbb{R}^n$.

On vient de définir un champ vectoriel « dans un esprit analytique», i.e. dans l'espace $\mathbb{R}^n$ à coordonnées $x_1, \ldots, x_n$ fixes. S'agissant de $\mathcal{A}$ affine (ou euclidien) quelconque, on exigerait que, dans un changement de coordonnées, les fonctions X_i varient en chaque point selon la loi vectorielle. Nous nous abstiendrons de considérer les champs vectoriels dans $\mathcal{A}$. Le fait est qu'ils présentent un défaut conceptuel (que nous ignorons pour le moment), et c'est seulement au cours du troisième semestre que nous serons à mesure de les définir « de façon correcte».

Nous nous permettons néanmoins d'écrire la formule intuitive

$$X = X_1\mathbf{e}_1 + \ldots + X_n\mathbf{e}_n \tag{6}$$

en comprenant par $\mathbf{e}_1, \ldots, \mathbf{e}_n$ la base canonique $(1, 0, \ldots, 0), \ldots$ $\ldots, (0, \ldots, 0, 1)$ de l'espace $\mathbb{R}^n$.

Avec ces notations, on a en particulier

$$\operatorname{grad} F = \frac{\partial F}{\partial x_1}\mathbf{e}_1 + \ldots + \frac{\partial F}{\partial x_n}\mathbf{e}_n.$$

* * *

Définition 3. Un point $x_0 \in U$ est un *point singulier* d'un champ vectoriel (différentiable) X si $X_i(\mathbf{x}_0) = 0$ pour tout $i = 1, \ldots, n$, i.e. si $X(\mathbf{x}_0) = 0$.

On souligne que la propriété d'un champ d'être différentiable est conservée en tout point singulier.

Ainsi, *l'ensemble de points $x \in U$ tels que $F(\mathbf{x}) = 0$, F étant une fonction différentiable, est une hypersurface différentiable s'il ne contient aucun point singulier du champ* grad F. □

D'ailleurs, le nom d'hypersurface reste attaché à cet ensemble même s'il renferme des points singuliers à condition que ceux-ci « ne soient pas trop nombreux » (sinon on aboutirait à un fermé quelconque par le théorème de Whitney). On suppose d'ordinaire que les points singuliers (à propos, ils s'appellent les *points singuliers de l'hypersurface $F = 0$*) sont isolés ou remplissent au moins une ou plusieurs « surfaces de dimension inférieure ».

Exemple. Le gradient de la forme quadratique

$$F(\mathbf{x}) = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2, \qquad \lambda_1 \neq 0, \ldots, \lambda_n \neq 0,$$

est représenté par la formule

$$\operatorname{grad} F = (2\lambda_1 x_1, \ldots, 2\lambda_n x_n)$$

et admet un point singulier $(0, \ldots, 0)$ unique. Aussi *l'hypersurface non dégénérée du second degré*

$$\lambda_1 x_1^2 + \ldots + \lambda_n x_n^2 = 1$$

(qui est un ellipsoïde ou un hyperboloïde) ne présente pas de points singuliers, i.e. c'est une hypersurface différentiable au sens de la définition 1.

Le cône du second degré

$$\lambda_1 x_1^2 + \ldots + \lambda_n x_n^2 = 0$$

possède par contre un point singulier unique qui est le sommet $(0, \ldots, 0)$.

Les points singuliers du cylindre sur le cône

$$\lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 = 0, \qquad \lambda_1 \neq 0, \ldots, \lambda_r \neq 0,$$

remplissent avec eux le plan $x_1 = 0, \ldots, x_r = 0$ de dimension $n - r$.

* * *

On définit pour les champs vectoriels une opération d'addition

$$(X + Y)_i (\mathbf{x}) = X_i (\mathbf{x}) + Y_i (\mathbf{x}), \quad i = 1, \ldots, n,$$

et une opération de multiplication par des fonctions

$$(fX)_i (\mathbf{x}) = f (\mathbf{x}) X_i (\mathbf{x}), \quad i = 1, \ldots, n.$$

On vérifie automatiquement que *l'ensemble* $\mathcal{X}$ (U) *de tous les champs vectoriels différentiables sur* U *est* pour ces opérations *un module sur l'anneau* $\mathcal{F}$ (U). □

Il nous paraît opportun d'introduire une définition algébrique générale.

Soient Λ un anneau quelconque et $\mathfrak{M}$ un module sur Λ. Une famille $m_1, \ldots, m_n$ d'éléments de $\mathfrak{M}$ en constitue une *base* si l'on associe à tout $m \in \mathfrak{M}$ des éléments bien définis $\lambda_1, \ldots, \lambda_n$ de Λ tels que

$$m = \lambda_1 m_1 + \ldots + \lambda_n m_n.$$

Contrairement aux espaces vectoriels (qui sont des modules sur un corps), *tout* module sur l'anneau Λ n'a pas de base. Un module ayant une base est dit *libre*. Si un module libre $\mathfrak{M}$ a toutes ses bases constituées d'un même nombre n d'éléments, on dit que le *module* $\mathfrak{M}$ *possède un rang* et que *ce rang est égal à* n. On trouve en général des anneaux tels que des modules libres sur ces anneaux n'ont pas de rang, mais ceux-ci sont nécessairement non commutatifs (essayer de démontrer !). Aussi *tout module libre sur l'anneau* $\mathcal{F}$ (U) *possède* en particulier *un rang*. □

Chaque vecteur $\mathbf{e}_i$ de la formule (6) s'interprète comme étant un champ vectoriel qui a toutes ses composantes $\equiv 0$ à part la i-ième qui est identiquement 1. La formule (6) signifie dans ce cas que *les champs* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *forment une base du module* $\mathcal{X}$ (U), ce qui démontre qu'*étant donné un ouvert* $U \subset \mathbb{R}^n$ *quelconque, le module* $\mathcal{X}$ (U) *des champs vectoriels sur* U *est un module libre de rang* n *sur l'anneau* $\mathcal{F}$ (U). □

D'autre part, le module $\mathcal{X}$ (U) est évidemment, à l'instar de l'anneau $\mathcal{F}$ (U), un espace vectoriel (de dimension infinie) sur $\mathbb{R}^n$.

* * *

L'application $F \to \operatorname{grad} F$ de l'anneau $\mathcal{F}$ (U) dans le module $\mathcal{X}$ (U) change, on le constate facilement, toute somme en une somme et tout produit par un nombre en un produit par un nombre, i.e. *c'est une application linéaire* (un homomorphisme) de l'espace vectoriel $\mathcal{F}$ (U) dans l'espace vectoriel $\mathcal{X}$ (U). Appliquée au produit de fonctions, elle opère par la formule

$$\operatorname{grad} : (FG) = F \operatorname{grad} G + G \operatorname{grad} F \tag{7}$$

(ce résultat provient de suite de la formule de dérivation des produits).

Le noyau de l'application linéaire

$$\text{grad}: F \longmapsto \text{grad } F$$

est évidemment *constitué de fonctions constantes localement*, i.e. de fonctions qui sont constantes sur chaque composante connexe de l'ensemble U.

L'image de l'application grad ne coïncide en général pas avec $\mathscr{X}(U)$.

Définition 4. Un champ vectoriel de la forme grad F s'appelle *champ de gradient* ou *champ de potentiel*. Si $X = \text{grad } F$, la fonction F est le *potentiel* du champ X. Le potentiel (quand il existe) est défini de façon unique à une fonction constante localement près.

Si

$$\frac{\partial X_i}{\partial x_j} = \frac{\partial X_j}{\partial x_i} \tag{8}$$

pour tout i et tout j ($i, j = 1, \ldots, n$) partout dans U, le champ (6) est dit *irrotationnel*. On vérifie de suite qu'*un champ de potentiel est nécessairement irrotationnel*. En effet, $\frac{\partial X_i}{\partial x_j} = \frac{\partial^2 F}{\partial x_i \partial x_j}$ et $\frac{\partial X_j}{\partial x_i} = \frac{\partial^2 F}{\partial x_j \partial x_i}$ si $X_i = \frac{\partial F}{\partial x_i}$. Or, la propriété connue des dérivées partielles mixtes entraîne $\frac{\partial^2 F}{\partial x_i \partial x_j} = \frac{\partial^2 F}{\partial x_j \partial x_i}$. □

En Analyse, on substitue souvent au champ (6) l'expression différentielle $X_1\, dx_1 + \ldots + X_n\, dx_n$ qui devient la *différentielle totale* dF d'une fonction F sous la condition nécessaire (8). Nous reverrons ce problème au cours du troisième semestre.

Les conditions (8) ne sont en général pas suffisantes, i.e. tout champ irrotationnel n'est pas de potentiel. Il l'est seulement pour U simple tel l'intérieur d'une boule ou d'un cube. Quant aux domaines $U \subset \mathbb{R}^n$ arbitraires, leur complexité se mesure à la dimension de l'espace quotient de l'espace vectoriel des champs irrotationnels par l'espace vectoriel des champs de potentiel. (Pour les détails, on attendra le troisième semestre.)

* * *

Dans le cas $n = 3$, les champs irrotationnels sont susceptibles d'une description plus facile. Nous supposerons donc partout dans la suite de la leçon que $n = 3$, et nous désignerons traditionnellement par $\mathbf{u}$, $\mathbf{v}$, ... les champs vectoriels de $\mathbb{R}^3$. Convenons des notations suivantes: P, Q, R, composantes d'un champ $\mathbf{u}$ (ce sont les notations d'usage); x, y, z, coordonnées x_1, x_2, x_3 dans $\mathbb{R}^3$; $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, vecteurs de base unitaires $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$; $\mathbf{r}$, vecteur $x\mathbf{i} + y\mathbf{j} +$

$+ z\mathbf{k}$. Le symbole U désignera une fois de plus un ouvert $U \subset \mathbb{R}^3$ dans lequel tous nos champs et toutes nos fonctions sont définis (et différentiables).

On note qu'avec ces notations,

$$\text{grad } F = \frac{\partial F}{\partial x}\mathbf{i} + \frac{\partial F}{\partial y}\mathbf{j} + \frac{\partial F}{\partial z}\mathbf{k}. \tag{9}$$

Définition 5. On appelle *rotationnel* du champ vectoriel

$$\mathbf{u} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k},$$

et on note rot $\mathbf{u}$, le champ vectoriel défini par

$$\text{rot } \mathbf{u} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k} \tag{10}$$

(la notation curl $\mathbf{u}$ très répandue naguère est presque abandonnée).

Il est clair que l'application

$$\text{rot} : \mathcal{X}(U) \to \mathcal{X}(U)$$

est un homomorphisme (un opérateur linéaire). Les éléments de son noyau ne sont autres que les champs irrotationnels, et dire que tout champ de potentiel est irrotationnel, c'est dire que

$$\text{rot grad } F = 0 \tag{11}$$

pour toute fonction $F \in \mathcal{F}(U)$.

Exemple 1. Un champ de la forme

$$\mathbf{u} = f(r)\,\mathbf{r},$$

où $r = |\mathbf{r}|$ et f est une fonction quelconque différentiable (pour $r > 0$), est dit *central*. Il est défini et différentiable partout, sauf au point (0, 0, 0). On a pour ce champ

$$P = f(r)\,x, \quad Q = f(r)\,y, \quad R = f(r)\,z.$$

D'autre part, on a de suite en dérivant $r = \sqrt{x^2 + y^2 + z^2}$:

$$\frac{\partial r}{\partial x} = \frac{x}{r}, \quad \frac{\partial r}{\partial y} = \frac{y}{r}, \quad \frac{\partial r}{\partial z} = \frac{z}{r}.$$

Par conséquent,

$$\begin{aligned}
&\frac{\partial P}{\partial x} = f'(r)\frac{x^2}{r} + f(r), \quad \frac{\partial P}{\partial y} = f'(r)\frac{xy}{r}, \quad \frac{\partial P}{\partial z} = f'(r)\frac{xz}{r},\\
&\frac{\partial Q}{\partial x} = f'(r)\frac{xy}{r}, \quad \frac{\partial Q}{\partial y} = f'(r)\frac{y^2}{r} + f(r), \quad \frac{\partial Q}{\partial z} = f'(r)\frac{yz}{r},\\
&\frac{\partial R}{\partial x} = f'(r)\frac{xz}{r}, \quad \frac{\partial R}{\partial y} = f'(r)\frac{yz}{r}, \quad \frac{\partial R}{\partial z} = f'(r)\frac{z^2}{r} + f(r).
\end{aligned} \tag{12}$$

Aussi on a, en particulier,

$$\frac{\partial R}{\partial y} = \frac{\partial Q}{\partial z}, \quad \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}, \quad \frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y},$$

i.e. rot $\mathbf{u} = 0$. Ainsi, *tout champ central est irrotationnel.* Il y a plus. *Chaque champ central dérive d'un potentiel.* En effet, on pose

$$F(\mathbf{r}) = \int_1^r rf(r)\,dr,$$

il vient de suite $\mathbf{u} = \operatorname{grad} F$. □

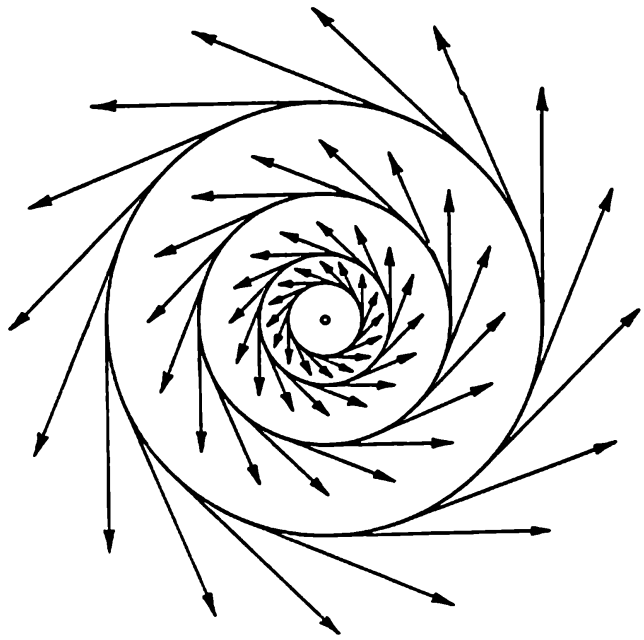
Si, en particulier, $f(r) = \frac{1}{r^3}$ et, partant, $|\mathbf{u}| = \frac{1}{r^2}$ (champ de pesanteur d'un point matériel), alors on a $F(\mathbf{r}) = -\frac{1}{r}$ à une constante près (*potentiel newtonien*).

Exemple 2. Soit

$$P = -y, \quad Q = x, \quad R = 0$$

(champ de vitesses d'une rotation plane), auquel cas

$$\operatorname{rot} \mathbf{u} = 2\mathbf{k}.$$



Champ de vitesses d'une rotation plane

L'opérateur rot agit sur le produit d'un champ vectoriel par une fonction suivant la formule

$$\operatorname{rot}(F\mathbf{u}) = F \operatorname{rot} \mathbf{u} + \operatorname{grad} F \times \mathbf{u} \tag{13}$$

qui est vérifiée de façon directe.

Le produit vectoriel de deux champs de (13) est naturellement considéré comme champ obtenu par multiplication vectorielle en chaque point des valeurs des facteurs.

Un champ de rotationnel (qui est donc de la forme rot **u**) est dit *solénoïdal*. Si **v** = rot **u**, le champ **u** s'appelle le *potentiel-vecteur* de **v**. Il est défini de façon unique à un champ irrotationnel près, i.e. à un terme de la forme grad F pour U simple.

* * *

Définition 6. On appelle la *divergence* du champ vectoriel

$$\mathbf{u} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k},$$

et on note div **u**, la fonction

$$\text{div } \mathbf{u} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}. \tag{14}$$

Si div **u** est identiquement nul, alors **u** est un *champ sans source*.

Exemple 3. S'agissant du champ central $\mathbf{u} = f(r)\,\mathbf{r}$, on a

$$\text{div } \mathbf{u} = 3f(r) + rf'(r) \tag{15}$$

(voir (12)). Si $f(r) = \frac{1}{r^3}$, alors div **u** = 0. Ainsi, *le champ de forces dérivant du potentiel newtonien est sans source.*

Il est immédiat d'établir que

$$\text{div rot } \mathbf{u} = 0$$

pour tout $\mathbf{u} \in \mathcal{X}(U)$. Ainsi, *un champ solénoïdal est nécessairement un champ sans source.* □

Comme plus haut, la réciproque n'est vraie que pour des domaines suffisamment « simples», et la dimension de l'espace quotient de l'espace vectoriel des champs sans source par l'espace vectoriel des champs solénoïdaux détermine le degré de complexité de U concerné.

L'application

$$\text{div} : \mathcal{X}(U) \to \mathcal{F}(U)$$

est manifestement linéaire, et on vérifie de suite que toute fonction $F \in \mathcal{F}(U)$ et tout champ $\mathbf{u} \in \mathcal{X}(U)$ satisfont à

$$\text{div } (F\mathbf{u}) = F \text{ div } \mathbf{u} + \mathbf{u} \text{ grad } F. \tag{16}$$

Nous avons donc défini trois applications linéaires

$$\mathcal{F}(U) \xrightarrow{\text{grad}} \mathcal{X}(U) \xrightarrow{\text{rot}} \mathcal{X}(U) \xrightarrow{\text{div}} \mathcal{F}(U)$$

jouissant des propriétés (7), (13), (16) respectivement et telles que les composées rot ∘ grad et div ∘ rot soient nulles.

* * *

La théorie de ces applications linéaires constitue l'Analyse vectorielle. Elle est d'un intérêt capital pour l'électromagnétisme théorique.

Chaque champ électromagnétique (par exemple, la lumière) est défini en chaque point du milieu par deux vecteurs, le vecteur intensité du champ électrique **E** et le vecteur intensité du champ magnétique **H**. Ces vecteurs sont des fonctions de points et de temps t, et on les définit complètement connaissant la densité de charges électriques ρ et la densité de courant **j**. Les grandeurs **E**, **H**, ρ et **j** sont reliées par les *équations de Maxwell* qui s'écrivent (dans un système d'unités correspondant)

$$\operatorname{div} \mathbf{E} = 4\pi\rho, \qquad \operatorname{div} \mathbf{H} = 0,$$

$$\operatorname{rot} \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t}, \qquad \operatorname{rot} \mathbf{H} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c}\mathbf{j},$$

c étant la vitesse de la lumière. Les équations de Maxwell constituent l'assise de la théorie de l'électromagnétisme, et elles sont en particulier la clef de voûte de l'optique et de la radioélectricité. C'est en fait pour permettre leur étude plus approfondie que l'Analyse vectorielle est née. Aujourd'hui, elle rend de grands services dans d'autres branches scientifiques (disons, en mécanique des milieux continus) sans parler de son intérêt mathématique certain.

Les chapitres les plus captivants de l'Analyse vectorielle relèvent des *formules intégrales* (dont nous nous occuperons pendant le troisième semestre). En attendant, nous nous bornerons à des formules simples qui ne contiennent pas les intégrales.

* * *

Il y a intérêt à établir ces formules à l'aide du *vecteur-champ symbolique de Hamilton*

$$\nabla = \frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}.$$

Si l'on prend pour produit, disons, du symbole $\frac{\partial}{\partial x}$ par une fonction P la dérivée partielle $\frac{\partial P}{\partial x}$, on considère le second membre de la formule (14) pour div **u** comme étant le produit scalaire du champ ∇ par le champ **u**. Ainsi,

$$\operatorname{div} \mathbf{u} = \nabla\mathbf{u}.$$

On représente de même rot **u** par le produit vectoriel

$$\operatorname{rot} \mathbf{u} = \nabla \times \mathbf{u}.$$

Soit dit en passant, cela permet d'exprimer rot **u** de façon élégante par le déterminant

$$\operatorname{rot} \mathbf{u} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & R \end{vmatrix}.$$

Quant au champ grad F, on l'écrit comme produit de ∇ par F:

$$\operatorname{grad} F = \nabla F$$

en mettant le facteur numérique après le vecteur.

* * *

Soient $\mathfrak{a}$ et $\mathfrak{b}$ qui sont, ou bien des fonctions, ou bien des champs vectoriels. On définit pour ces objets mathématiques plusieurs opérations de multiplication (si $\mathfrak{a}$ et $\mathfrak{b}$ sont par exemple des champs vectoriels, on les multiplie soit scalairement, soit vectoriellement). Soient $\underset{1}{*}$ et $\underset{2}{*}$ deux opérations de multiplication telles que le produit $\Delta \underset{1}{*} (\mathfrak{a} \underset{2}{*} \mathfrak{b})$ ait un sens.

La règle usuelle de dérivation d'un produit dit que la dérivée du produit de deux termes est égale à la somme des produits de chacun des termes par la dérivée de l'autre. Il est clair qu'on procède de même pour l'« opération » ∇. Aussi

$$\nabla \underset{1}{*} (\mathfrak{a} \underset{2}{*} \mathfrak{b}) = \nabla \underset{1}{*} (\mathfrak{a} \underset{2}{*} \overset{\downarrow}{\mathfrak{b}}) + \nabla \underset{1}{*} (\overset{\downarrow}{\mathfrak{a}} \underset{2}{*} \mathfrak{b}), \tag{17}$$

la flèche verticale repérant le terme soumis à l'action de ∇.

Supposons que $\mathfrak{a}$ et $\mathfrak{b}$ sont des fonctions F et G respectivement (si bien que l'opération $\underset{2}{*}$ est une multiplication des fonctions et $\underset{1}{*}$ une multiplication d'un champ vectoriel par une fonction). On a

$$\nabla (FG) = \nabla (F\overset{\downarrow}{G}) + \nabla (\overset{\downarrow}{F}G).$$

Mais il est évident que $\nabla (F\overset{\downarrow}{G}) = F (\nabla G)$ et, de même, $\nabla (\overset{\downarrow}{F}G) = G (\nabla F)$. Par conséquent,

$$\nabla (FG) = F (\nabla G) + G (\nabla F).$$

Ce n'est rien autre que la formule (7).

Si $\mathfrak{a}$ est une fonction F et $\mathfrak{b}$ un champ **u**, alors (17) donne lieu à deux formules

$$\nabla (F\mathbf{u}) = \nabla (F\overset{\downarrow}{\mathbf{u}}) + \nabla (\overset{\downarrow}{F}\mathbf{u})$$

et

$$\nabla \times (F\mathbf{u}) = \nabla \times (F\overset{\downarrow}{\mathbf{u}}) + \nabla \times (\overset{\downarrow}{F}\mathbf{u}).$$

Dans la première, $\nabla (\overset{\downarrow}{F\mathbf{u}}) = F (\nabla \mathbf{u})$ et $\nabla (\overset{\downarrow}{F\mathbf{u}}) = (\nabla F)\, \mathbf{u}$, de sorte que

$$\nabla (F\mathbf{u}) = F (\nabla \mathbf{u}) + (\nabla F)\, \mathbf{u}.$$

C'est (16).

On a de même $\nabla \times (\overset{\downarrow}{F\mathbf{u}}) = F (\nabla \times \mathbf{u})$ et $\nabla \times (\overset{\downarrow}{F\mathbf{u}}) = (\nabla F) \times \mathbf{u}$ dans la seconde formule obtenue à partir de (17). Aussi

$$\nabla \times (F\mathbf{u}) = F (\nabla \times \mathbf{u}) + \nabla F \times \mathbf{u}.$$

C'est l'égalité (13).

Plaçons-nous enfin dans le cas où $\mathfrak{a}$ et $\mathfrak{b}$ sont des champs $\mathbf{u}$ et $\mathbf{v}$, il vient trois formules nouvelles

$$\nabla (\mathbf{uv}) = \nabla (\overset{\downarrow}{\mathbf{u}}\mathbf{v}) + \nabla (\overset{\downarrow}{\mathbf{u}}\mathbf{v}), \tag{17a}$$

$$\nabla (\mathbf{u} \times \mathbf{v}) = \nabla (\mathbf{u} \times \overset{\downarrow}{\mathbf{v}}) + \nabla (\overset{\downarrow}{\mathbf{u}} \times \mathbf{v}), \tag{17b}$$

$$\nabla \times (\mathbf{u} \times \mathbf{v}) = \nabla \times (\mathbf{u} \times \overset{\downarrow}{\mathbf{v}}) + \nabla \times (\overset{\downarrow}{\mathbf{u}} \times \mathbf{v}). \tag{17c}$$

C'est la deuxième formule qu'on explique le plus facilement. En effet, on utilise les propriétés du produit mixte et on a de suite

$$\nabla (\mathbf{u} \times \overset{\downarrow}{\mathbf{v}}) = \nabla \mathbf{u}\overset{\downarrow}{\mathbf{v}} = -\mathbf{u}\nabla\overset{\downarrow}{\mathbf{v}} = -\mathbf{u}\nabla\mathbf{v} = -\mathbf{u}\,(\nabla \times \mathbf{v})$$

et

$$\nabla (\overset{\downarrow}{\mathbf{u}} \times \mathbf{v}) = \nabla \overset{\downarrow}{\mathbf{u}}\mathbf{v} = \mathbf{v}\nabla\overset{\downarrow}{\mathbf{u}} = \mathbf{v}\nabla\mathbf{u} = \mathbf{v}\,(\nabla \times \mathbf{u}).$$

La formule (17b) équivaut donc à

$$\operatorname{div} (\mathbf{u} \times \mathbf{v}) = (\operatorname{rot} \mathbf{u})\, \mathbf{v} - \mathbf{u} \operatorname{rot} \mathbf{v}. \tag{18}$$

On conçoit que ce « donc » est on ne peut plus conventionnel. En effet, nous n'avons pas justifié le fait d'avoir muni les produits contenant ∇ des propriétés du produit mixte. En plus que cette justification nous entraînerait trop loin, elle exigerait qu'on légitime plus sérieusement la formule de départ (17) que nous avons, à vrai dire, acceptée sans démonstration. Aussi tout ce que nous avons dit plus haut n'est rien d'autre que des considérations mneumoniques ou, au mieux, heuristiques qui permettent de réunir en (17) les formules (7), (16), (13) et (18) primitivement autonomes. Quant à la démonstration formelle de celles-ci (et, en particulier, de la formule nouvelle (18)), il ne nous reste qu'à les vérifier directement une à une.

Les formules (7), (16), (13), (18) n'épuisent pas (17) : les formules symboliques (17a) et (17c) sont encore à déchiffrer. A cet effet,

nous avons besoin du

Lemme. *On a pour trois vecteurs* **a**, **b**, **c** *quelconques*

$$\mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = (\mathbf{cb})\,\mathbf{a} - (\mathbf{ac})\,\mathbf{b}. \tag{19}$$

Démonstration. Prenons un repère orthonormé **i**, **j**, **k** tel que **a** soit colinéaire à **i** et que **b** soit coplanaire avec **i** et **j**. Alors

$$\begin{aligned} \mathbf{a} &= a_1\mathbf{i}, \\ \mathbf{b} &= b_1\mathbf{i} + b_2\mathbf{j}, \\ \mathbf{c} &= c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}, \end{aligned}$$

donc

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= (a_1b_2)\,\mathbf{k}, \\ \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) &= (a_1b_2c_2)\,\mathbf{i} - (a_1b_2c_1)\,\mathbf{j}. \end{aligned}$$

D'autre part,

$$\mathbf{cb} = c_1b_1 + c_2b_2, \qquad \mathbf{ac} = a_1c_1,$$

si bien que

$$\begin{aligned} (\mathbf{cb})\,\mathbf{a} - (\mathbf{ac})\,\mathbf{b} &= (c_1b_1 + c_2b_2)\,a_1\mathbf{i} - a_1c_1\,(b_1\mathbf{i} + b_2\mathbf{j}) = \\ &= (c_2b_2a_1)\,\mathbf{i} - (a_1c_1b_2)\,\mathbf{j}. \end{aligned}$$

Par conséquent, $\mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = (\mathbf{cb})\,\mathbf{a} - (\mathbf{ac})\,\mathbf{b}$. □

On se servira du lemme si l'un des facteurs est le champ vectoriel ∇, i.e., cette fois encore, dans des buts purement mneumoniques et heuristiques. Les formules obtenues seront correctes si nous donnons en outre une interprétation à $\mathbf{a}\nabla$, avec $\mathbf{a} = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$ un champ vectoriel, autre que $\nabla\mathbf{a} = \operatorname{div}\mathbf{a}$ qui s'impose, i.e. si nous renonçons pour le champ symbolique ∇ à la commutativité de la multiplication scalaire.

Plus précisément, nous considérerons $\mathbf{a}\nabla$ comme opérateur agissant sur le champ vectoriel $\mathbf{u} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ selon la formule

$$(\mathbf{a}\nabla)\,\mathbf{u} = A\frac{\partial P}{\partial x} + B\,\frac{\partial Q}{\partial y} + C\,\frac{\partial R}{\partial z}.$$

Avec cette convention, nous obtenons à l'aide de (19)

$$\nabla \times (\mathbf{u} \times \overset{\downarrow}{\mathbf{v}}) = (\nabla\mathbf{v})\,\mathbf{u} - (\mathbf{u}\nabla)\mathbf{v} = (\operatorname{div}\mathbf{v})\,\mathbf{u} - (\mathbf{u}\nabla)\,\mathbf{v},$$

$$\nabla \times (\overset{\downarrow}{\mathbf{u}} \times \mathbf{v}) = (\mathbf{v}\nabla)\,\mathbf{u} - (\nabla\mathbf{u})\,\mathbf{v} = (\mathbf{v}\nabla)\,\mathbf{u} - (\operatorname{div}\mathbf{u})\,\mathbf{v},$$

et la formule (17c) aboutit à

$$\operatorname{rot}(\mathbf{u} \times \mathbf{v}) = (\mathbf{v}\nabla)\,\mathbf{u} - (\mathbf{u}\nabla)\,\mathbf{v} + (\operatorname{div}\mathbf{v})\,\mathbf{u} - (\operatorname{div}\mathbf{u})\,\mathbf{v}. \tag{20}$$

On traite de même (17a) en récrivant (19)

$$\mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a}\,(\mathbf{cb}) - (\mathbf{ca})\,\mathbf{b},$$

il vient

$$\mathbf{u} \times (\nabla \times \mathbf{v}) = \nabla\,(\overset{\downarrow}{\mathbf{u}\mathbf{v}}) - (\mathbf{u}\nabla)\,\mathbf{v}$$

et

$$\mathbf{v} \times (\nabla \times \mathbf{u}) = \nabla (\overset{\downarrow}{\mathbf{u}}\mathbf{v}) - (\mathbf{v}\nabla)\,\mathbf{u},$$

donc,

$$\nabla (\overset{\downarrow}{\mathbf{u}}\mathbf{v}) + \nabla (\overset{\downarrow}{\mathbf{u}}\mathbf{v}) = \mathbf{u} \times (\nabla \times \mathbf{v}) + \mathbf{v} \times (\nabla \times \mathbf{u}) + (\mathbf{v}\nabla)\,\mathbf{u} + (\mathbf{u}\nabla)\,\mathbf{v}.$$

Ainsi, la formule (17a) fournit

(21) $\quad \operatorname{grad}(\mathbf{uv}) = \mathbf{u} \times \operatorname{rot}\mathbf{v} + \mathbf{v} \times \operatorname{rot}\mathbf{u} + (\mathbf{v}\nabla)\,\mathbf{u} + (\mathbf{u}\nabla)\,\mathbf{v}.$

Comme pour (7), (16), (13), (18), on démontre formellement les formules (20) et (21) par des calculs directs.

* * *

Les composantes des opérateurs grad, rot et div sont reliées elles aussi par des relations intéressantes.

On sait que

$$\operatorname{rot} \circ \operatorname{grad} = 0,$$
$$\operatorname{div} \circ \operatorname{rot} = 0.$$

On observe que l'hamiltonien réduit ces formules à l'affirmation « le produit vectoriel de deux vecteurs identiques est nul :

$$\operatorname{rot}(\operatorname{grad} F) = \nabla \times (\nabla F) = (\nabla \times \nabla)\,F = 0$$

et

$$\operatorname{div}(\operatorname{rot}\mathbf{u}) = \nabla (\nabla \times \mathbf{u}) = (\nabla \times \nabla)\,\mathbf{u} = 0 \text{»}.$$

Un rôle particulièrement important incombe à l'opérateur

$$\Delta = \operatorname{div} \circ \operatorname{grad}$$

qu'on considère comme étant le carré scalaire ∇^2 de l'hamiltonien. On l'appelle l'*opérateur de Laplace* (ou le laplacien). Il opère de $\mathcal{F}(U)$ dans $\mathcal{F}(U)$ conformément à la formule

$$\Delta F = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + \frac{\partial^2 F}{\partial z^2}.$$

On écrit en termes de laplacien les équations fondamentales de la physique mathématique (qui font l'objet d'un cours universitaire spécial).

Une fonction F est dite *harmonique* si $\Delta F = 0$. Un exemple en est le potentiel newtonien $F = -\frac{1}{r}$ (voir plus haut). Le cours d'équations de la physique mathématique enseigne que *toute fonction harmonique est un potentiel du champ gravitationnel d'une masse.* Ce résultat à lui seul suffit pour situer les fonctions harmoniques en physique (donc en mathématiques).

On applique l'opérateur Δ aux champs vectoriels en le faisant agir sur chaque composante prise séparément : si $\mathbf{u} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$, alors

$$\Delta\mathbf{u} = (\Delta P)\,\mathbf{i} + (\Delta Q)\,\mathbf{j} + (\Delta R)\,\mathbf{k},$$

auquel cas

$$\mathrm{rot} \circ \mathrm{rot} = \mathrm{grad} \circ \mathrm{div} - \Delta.$$

En effet, le lemme implique

$$\mathrm{rot}\,(\mathrm{rot}\,\mathbf{u}) = \nabla \times (\nabla \times \mathbf{u}) = \nabla\,(\nabla\mathbf{u}) - (\nabla\nabla)\,\mathbf{u}. \quad \square$$

On écrit d'autres expressions différentielles. Etant données deux fonctions F et G quelconques, on définit par exemple le produit scalaire de leurs gradients :

$$\Delta\,(F, G) = \mathrm{grad}\ F\ \mathrm{grad}\ G = \frac{\partial F}{\partial x}\frac{\partial G}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial G}{\partial y} + \frac{\partial F}{\partial z}\frac{\partial G}{\partial z}$$

qui porte le nom de *paramètre différentiel mixte de Beltrami* des fonctions F et G. On a en particulier pour $F = G$

$$\Delta_2 F = (\mathrm{grad}\ F)^2 = \left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2 + \left(\frac{\partial F}{\partial z}\right)^2.$$

Ce carré scalaire du gradient s'appelle le *premier paramètre différentiel de Beltrami* de la fonction F.

Le produit mixte des gradients de trois fonctions est par définition le *paramètre différentiel de Darboux*. Il n'est rien d'autre que le *jacobien* de la transformation définie par trois fonctions données, ce qui explique l'emploi très restreint du terme.

LEÇON 23

Courbes continues, différentiables et régulières. — Courbes équivalentes. — Courbes régulières dans le plan et graphes des fonctions. — Hyperplan tangent à une hypersurface. — Longueur d'une courbe. — Courbes dans le plan. — Courbes dans l'espace de dimension 3.

Si l'on définit la notion d'une courbe considérée comme trajectoire d'un point mobile, on aboutit à la

Définition 1. On appelle *courbe continue* dans l'espace euclidien (ou affine) $\mathscr{E}$ de dimension n l'application continue

$$\mathbf{x}: t \mapsto \mathbf{x}(t) \tag{1}$$

d'un segment $[a, b]$, $a < b$, de l'axe des t dans l'espace $\mathscr{E}$ (on sous-entend que les points de $\mathscr{E}$ sont caractérisés par leurs rayons vecteurs par rapport à un point O fixe).

La continuité des applications (1) a un sens parce que l'espace euclidien $\mathscr{E}$ est un espace métrique. On montre sans peine (le faire !) que cette propriété de (1) est équivalente à la continuité de n fonctions numériques

$$x_i: t \mapsto x_i(t), \quad i = 1, \ldots, n, \tag{2}$$

où $x_1(t), \ldots, x_n(t)$ sont les coordonnées du vecteur $\mathbf{x}(t)$ dans une base quelconque. Comme la base est à priori indépendante de la métrique (car elle n'est pas orthonormée), l'application (1) continue pour une métrique l'est dans toute autre. Ainsi, la propriété considérée de (1) est indépendante de la métrique, i.e. c'est une propriété affine. Autrement dit, les applications continues (1) ont également un sens dans le cas de $\mathscr{E}$ affine (i.e. dans un espace non métrique). Cf. I.12, déf.1.

Le lecteur a certes fait connaissance avec la définition 1 (en termes de $\mathbb{R}^n$) par le cours d'Analyse.

On souligne qu'une courbe continue est par définition non un ensemble de points, mais une application; n'empêche que le vocabulaire utilisé le contredit constamment. On dit par exemple que la courbe (1) *passe par le point* $\mathbf{x}_0$ s'il existe une valeur t_0 (qui n'est en général pas unique) du paramètre t telle que $\mathbf{x}(t_0) = \mathbf{x}_0$. Le point $\mathbf{x}(a)$ s'appelle le *point initial* de (1) et $\mathbf{x}(b)$ en est le *point final*. On dit également que la courbe (1) *joint* le point $\mathbf{x}(a)$ à $\mathbf{x}(b)$, et ainsi de suite.

On appelle des fois *support* de la courbe (1) l'ensemble de tous ses points, i.e. l'image de $[a, b]$ par l'application (1).

La définition 1 a été proposée au XIXe siècle par Camille Jordan. Jordan avait la conviction (partagée d'ailleurs par tous les mathématiciens de l'époque) que sa définition reflétait suffisamment bien l'idée intuitive qu'on a d'une courbe. Cette conviction n'a pas fait long feu. La construction par Peano d'une courbe continue qui passe (plusieurs fois en général) par chaque (sic !) point d'un carré a fait l'effet d'une bombe. Car on a compris que la condition de continuité à elle seule ne suffit pas.

Dans la leçon 22, nous avons défini une fonction F différentiable sur un ouvert $U \subset \mathbb{R}^n$. Considérons un ensemble $C \subset \mathbb{R}^n$ arbitraire et une fonction f définie sur C. Nous dirons que f est *différentiable sur* C s'il existe un ouvert $U \subset \mathbb{R}^n$ et une fonction F différentiable sur U tels que $C \subset U$ et

$$f = F|_C.$$

Etant donnée une fonction $t \mapsto x(t)$ définie sur $[a, b]$, nous l'appellerons en particulier *différentiable* s'il existe, sur un intervalle ouvert contenant $[a, b]$, une fonction différentiable qui coïncide avec $x(t)$ sur $[a, b]$.

Définition 2. L'application (1) s'appelle *courbe différentiable* dans $\mathscr{E}$ si les fonctions coordonnées (2) sont des fonctions différentiables sur $[a, b]$.

La définition est évidemment intrinsèque (i.e. indépendante du système de coordonnées).

On définit pour toute courbe (1) différentiable et tout $t \in [a, b]$ une limite

$$\mathbf{x}'(t) = \lim_{\Delta t \to 0} \frac{\mathbf{x}(t+\Delta t) - \mathbf{x}(t)}{\Delta t} \tag{3}$$

appelée *vecteur tangent* à la courbe (1) au point t (ou au point $\mathbf{x}(t)$). Ses coordonnées sont visiblement les dérivées

$$x_1'(t), \ldots, x_n'(t) \tag{4}$$

des coordonnées (2) du vecteur $\mathbf{x}(t)$. Au lieu de (3) on emploie également la notation $\frac{d\mathbf{x}(t)}{dt}$.

On répète le procédé autant de fois qu'on le veut et on obtient $\mathbf{x}''(t)$, $\mathbf{x}'''(t)$, . . ., dont les coordonnées sont les dérivées correspondantes des fonctions (2).

On constate sans peine que si une courbe continue (1) admet une limite (3), cela entraîne l'existence des dérivées (4). Ainsi, la condition de (1) différentiable équivaut à exiger l'existence des dérivées $\mathbf{x}'(t)$, $\mathbf{x}''(t)$, . . . de tout ordre (voulu). Le fait confirme l'indépendance de la condition de fonction ou de courbe différentiable vis-à-vis du système de coordonnées.

Les courbes différentiables (plus précisément leurs supports) répondent en gros à l'idée intuitive qu'on a d'une courbe. On montrera au cours du troisième semestre qu'une courbe différentiable ne passe en tout cas pas par tous les points d'un carré (et même que l'ensemble de tous ses points, i.e. le support de la courbe, est un « ensemble de mesure nulle»). D'autre part, une courbe différentiable peut ne pas être « lisse» du tout et présenter des points anguleux (comme, par exemple, la courbe plane $x = t^2$, $y = t^3$). Pour éviter ces cas pathologiques, on introduit la

Définition 3. Une courbe (1) différentiable est dite *régulière* si $\mathbf{x}'(t) \neq 0$ pour tout $t \in [a, b]$.

On verra que la notion de courbe régulière suffit complètement pour se représenter intuitivement une courbe « lisse». En attendant, on se propose d'examiner une question importante.

* * *

Géométriquement, les définitions 1 à 3 ont un défaut majeur, à savoir les « courbes» ainsi introduites ne sont pas des ensembles. D'autre part, définir une courbe comme image de $[a, b]$ par une application continue (différentiable ou régulière) de ce segment dans l'espace $\mathscr{E}$ est absolument incorrect pour plusieurs raisons. On donne une interprétation plus géométrique et plus intuitive d'une courbe (qui est de plus d'une exploitation fructueuse) avec la

Définition 4. Deux courbes

$$\mathbf{x}: t \longmapsto \mathbf{x}(t), \quad \mathbf{x}_1: t_1 \longmapsto \mathbf{x}_1(t_1),$$

avec $a \leqslant t \leqslant b$ et $a_1 \leqslant t_1 \leqslant b_1$ respectivement, sont dites *équivalentes* s'il existe une fonction

$$\varphi: t \longmapsto \varphi(t) \tag{5}$$

telle que $\varphi(a) = a_1$, $\varphi(b) = b_1$ et $\mathbf{x}(t) = \mathbf{x}_1(\varphi(t))$ quel que soit $t \in [a, b]$. On dit de la fonction (5) que c'est un *changement* de paramètre sur t.

Il est clair que deux courbes équivalentes ont même support.

Les classes de courbes équivalentes s'appellent les *courbes non paramétrées*. Ce sont les *courbes* de certains auteurs (surtout des traditionalistes), tandis que les courbes au sens des définitions 1 à 3 sont pour eux les *courbes paramétrées* ou les *chemins*. Intuitivement, passer à une courbe équivalente c'est modifier la vitesse à laquelle le point parcourt la même trajectoire. On conçoit que le changement de vitesse ne peut être quelconque. Soit, par exemple, une courbe continue. On demande en principe que la fonction φ soit un homéomorphisme (i.e. une bijection bicontinue) du segment $[a, b]$ sur le segment $[a_1, b_1]$. Aûtrement dit, on impose à φ d'être continue et strictement monotone (auquel cas la fonction réciproque existe et est continue). Dans le cas contraire, la relation entre les courbes

au sens de la définition 4 n'est en général pas une relation d'équivalence sur l'ensemble de toutes les courbes, d'où l'impossibilité d'introduire les classes de courbes équivalentes. On peut cependant faire l'hypothèse de fonctions (5) non strictement monotones (si bien que les fonctions réciproques sont discontinues) pourvu que la courbe $t \mapsto x_1 (\varphi (t))$ reste continue pour φ discontinue. Cela signifie que le point mobile est autorisé à marquer un temps d'arrêt et que, inversement, un point stationnaire, puis animé d'un mouvement équivalent, brûle cette étape. Il y a plus. Avec une définition 4 plus compliquée, on suppose (5) non monotones quelconques (ce qui permet un mouvement rétrograde du point). Toutes ces questions sont étudiées en détail en Analyse. Fidèles à notre tactique générale, nous nous bornerons pour notre part *aux changements de paramètre réguliers*, i.e. à des fonctions (5) différentiables telles que

$$\varphi' (t) > 0 \text{ pour tout } t \in [a, b].$$

Cela garantit la condition de régularité dans les changements de paramètre.

Il ne faut pas surestimer l'intérêt de la notion de courbe non paramétrée car, *primo*, elle est d'une unité (d'une « équivalence») plus compliquée que la notion de courbe paramétrée, et, *secundo*, elle ne traduit quand même assez bien l'idée intuitive d'une courbe en tant qu'ensemble ponctuel (deux courbes ayant même support ne sont pas nécessairement équivalentes). Si, au début de notre siècle, on préférait nettement la notion de courbe non paramétrée qui serait la plus susceptible—d'une interprétation géométrique sensible, ces dernières années, on voit s'affermir la tendance à mettre au premier plan les courbes paramétrées. En effet, il s'agit d'un concept plus simple et, chose essentielle, ce sont justement les courbes paramétrées qui interviennent dans les constructions mathématiques réelles. Cela explique en particulier pourquoi le nom de courbe a été naguère donné aux courbes non paramétrées, tandis qu'aujourd'hui on le donne de plus en plus souvent aux courbes paramétrées.

On ne saurait certes oublier que de nombreuses notions et constructions naturelles et commodes ne sont pas conservées par passage aux courbes équivalentes, si bien qu'il est impossible de les définir pour les courbes non paramétrées. C'est par exemple le cas du vecteur tangent qui est dans ce passage multiplié par $\varphi' (t)$. Cette circonstance oblige jusqu'aux plus chauds partisans des courbes non paramétrées à les abandonner dans la pratique pour les courbes paramétrées tout en expliquant leur trahison par le caractère « naturel» du paramètre introduit (voir plus bas).

Aussi nous nous occuperons essentiellement des courbes paramétrées, et les courbes équivalentes ne nous intéresseront qu'au fur et à mesure.

* * *

Maintenant nous sommes à même d'établir à quel point la notion de courbe régulière concorde avec l'idée intuitive qu'on a d'une courbe. On se place dans le plan (ce qui simplifie l'exposé), et on désigne une fois de plus les coordonnées dans le plan par les lettres x, y.

Le graphe d'une fonction différentiable quelconque $y = y(x)$ est le support de la courbe régulière

$$x = t, \quad y = y(t),$$

que nous appellerons par abus de langage le graphe de $y(x)$.

Quelles courbes planes répondent à l'idée intuitive d'une courbe « lisse »? On exige visiblement que

a) le graphe de toute fonction différentiable (pour toute position des axes de coordonnées) soit une courbe « lisse »,

b) une courbe (régulièrement) équivalente à une courbe « lisse » soit une courbe « lisse »,

c) une courbe est « lisse » si et seulement si elle l'est localement, i.e. au voisinage de chacun de ses points.

La plus petite classe de courbes vérifiant ces conditions est formée de courbes *localement équivalentes* (i.e. équivalentes au voisinage de tous leurs points) aux graphes des fonctions différentiables (variant d'un point à l'autre). Il est clair que ces courbes sont régulières. La réciproque est également vraie (ce qui justifie du point de vue intuitif la formation de la classe des courbes régulières). Ainsi, *toute courbe plane régulière est localement équivalente au graphe d'une fonction différentiable.*

En effet, si la courbe

$$x = x(t), \quad y = y(t), \quad a \leqslant t \leqslant b, \tag{6}$$

est régulière, alors on a pour tout point $t_0 \in [a, b]$, ou bien $x'(t_0) \neq 0$, ou bien $y'(t_0) \neq 0$. On suppose, pour fixer les idées, que $x'(t_0) \neq 0$, auquel cas le théorème des fonctions implicites (appliqué à $F(x, t) = x - x(t)$) entraîne que $t \mapsto x(t)$ est *localement inversible*, i.e. il existe un voisinage U_0 de t_0 et un voisinage V_0 de $x_0 = x(t_0)$ tels que $t \mapsto x(t)$ établisse la bijection $U_0 \to V_0$, la fonction inverse $x \mapsto t(x)$ étant différentiable. Cela étant, $t'(x) \neq 0$ pour tout $x \in V_0$, de sorte que si $t'(x) > 0$ sur V_0, alors $x \mapsto x(t)$ est un changement régulier sur le paramètre pour (6) dans U_0. Dans cette application, la courbe (6) devient (dans le voisinage U_0) une courbe équivalente qui représente (sur V_0) la fonction différentiable $y = y(t(x))$. Dans le cas de $t'(x) < 0$ sur V_0, on prend pour nouveau paramètre $-x$ (i.e. on inverse la direction de l'axe des abscisses). Soit enfin $x'(t_0) = 0$. Le nouveau paramètre est ici y (ou $-y$). □

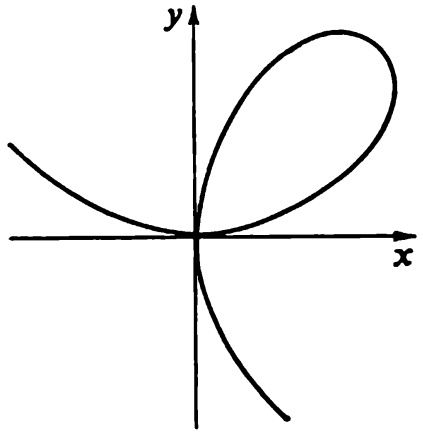
On note que la propriété locale de l'affirmation est établie « relativement à un paramètre », i.e. on considère la restriction de la courbe

à un voisinage du point $t_0 \in [a, b]$. S'agissant du voisinage du point $(x(t_0), y(t_0))$ du plan, l'affirmation analogue n'a pas de sens.

Exemple. La courbe

$$x = \frac{3t(1-t)^2}{3t^2-3t+1}, \qquad y = \frac{3t^2(1-t)}{3t^2-3t+1}$$

appelée *folium de Descartes* passe par le point (0, 0) pour $t = 0$ et pour $t = 1$. Au voisinage de $t = 0$, elle est équivalente au graphe d'une fonction $y = y(x)$, et au voisinage de $t = 1$ elle l'est à celui de $x = x(y)$. Dans le plan au voisinage de (0, 0) cette courbe (plus précisément, son support) est la réunion de deux graphes.



Folium de Descartes

Le point (0, 0) est le *nœud* du folium de Descartes, et deux graphes constitutifs de la courbe au voisinage du nœud en sont les *branches*. Nous ne nous proposons pas d'insister sur ces questions car nous nous bornerons dans la suite à l'étude locale des courbes sur des portions suffisamment petites de l'axe des t (i.e. au cas où les courbes sont équivalentes aux graphes). Aussi on observe seulement que si les courbes régulières (plus précisément, leurs supports) ne sont pas des hypersurfaces régulières dans le plan au sens de la définition 1 de la leçon précédente, la cause en est la présence de nœuds. D'ailleurs, nous ne pouvons pas affirmer pour le moment (si ce n'est pas pour le cas local) que le support de chaque courbe plane dépourvue de nœuds est une hypersurface régulière ni que, inversement, toute hypersurface régulière dans le plan est le support d'une courbe régulière (qui est forcément sans nœuds). Ces questions seront étudiées dans leur généralité naturelle pendant le troisième semestre.

* * *

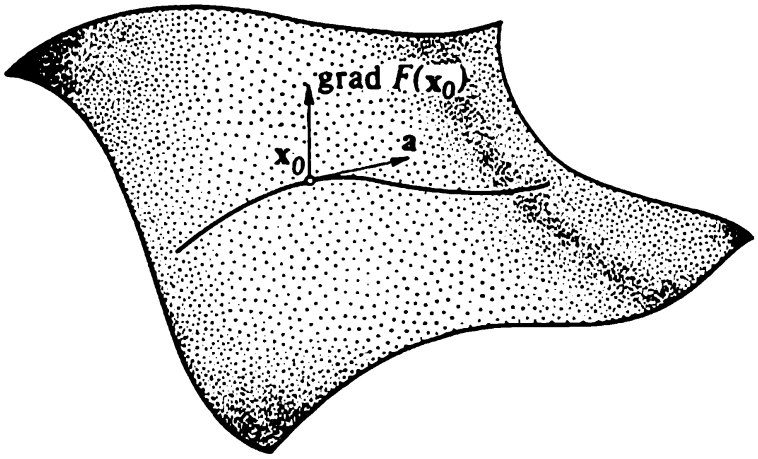
Conformément au vocabulaire de la théorie des courbes, nous dirons que la courbe (1) *est située* sur l'hypersurface

(7) $$F(\mathbf{x}) = 0$$

de l'espace $\mathscr{E}$ si $F(\mathbf{x}(t)) = 0$ quel que soit $t \in [a, b]$, i.e. si le support de la courbe appartient à (7).

Définition 5. Un vecteur $\mathbf{a}$ est dit *tangent* à l'hypersurface (7) en son point $\mathbf{x}_0$ si l'hypersurface contient une courbe $t \longmapsto \mathbf{x}(t)$ passant par $\mathbf{x}_0$ pour $t = t_0$ telle que $\mathbf{a}$ soit un vecteur tangent à cette courbe au point t_0, i.e. si $\mathbf{a} = \mathbf{x}'(t_0)$.

Soient $\mathscr{V}$ l'espace vectoriel associé à l'espace ponctuel (qu'on suppose euclidien pour fixer les idées) $\mathscr{E}$ et $\mathscr{H}_{\mathbf{x}_0}$ l'ensemble des vecteurs tangents à (7) (supposée régulière) en son point $\mathbf{x}_0$.



Vecteur tangent

Proposition 1. *L'ensemble $\mathscr{H}_{\mathbf{x}_0}$ est un sous-espace de dimension $n-1$ de $\mathscr{V}$ constitué de tous les vecteurs orthogonaux au vecteur* grad $F(\mathbf{x}_0)$:

$$\mathscr{H}_{\mathbf{x}_0} = \{\mathbf{a} \in \mathscr{V} \,;\, \mathbf{a} \text{ grad } F(\mathbf{x}_0) = 0\}.$$

Démonstration. Si $\mathbf{a} \in \mathscr{H}_{\mathbf{x}_0}$, alors il existe une courbe $t \longrightarrow \mathbf{x}(t)$, $a \leqslant t \leqslant b$, de $\mathscr{V}$ pour laquelle

(8) $$F(\mathbf{x}(t)) = 0 \text{ quels que soient } t \in [a, b]$$

et

(9) $$\mathbf{x}_0 = \mathbf{x}(t_0), \quad \mathbf{a} = \mathbf{x}'(t_0).$$

Mais la formule d'Analyse qui donne la dérivée d'une fonction composée

$$F(\mathbf{x}(t)) = F(x_1(t), \ldots, x_n(t))$$

se récrit

$$\frac{dF(\mathbf{x}(t))}{dt} = \mathbf{x}'(t) \text{ grad } F(\mathbf{x}(t))$$

(il est naturel que les coordonnées $x_1, \ldots, x_n$ sont supposées rectangulaires). On dérive (8) et on pose $t = t_0$, il vient par suite des

égalités (9)

$$\mathbf{a} \operatorname{grad} F(\mathbf{x}_0) = 0. \tag{10}$$

Inversement, soit (10). Sans restreindre la généralité, on choisit un système de coordonnées tel que grad $F(\mathbf{x}_0)$ soit parallèle à l'axe Ox_n, auquel cas on a

$$\frac{\partial F}{\partial x_1}(\mathbf{x}_0) = 0, \ldots, \quad \frac{\partial F}{\partial x_{n-1}}(\mathbf{x}_0) = 0, \quad \frac{\partial F}{\partial x_n}(\mathbf{x}_0) \neq 0, \tag{11}$$

et la condition (10) s'écrit $a_n = 0$. Comme $\frac{\partial F}{\partial x_n}(\mathbf{x}_0) \neq 0$, l'hypersurface (7) est au voisinage de $\mathbf{x}_0$ le graphe d'une fonction différentiable

$$x_n = \varphi(\hat{\mathbf{x}}).$$

En d'autres termes,

$$x_n^{(0)} = \varphi(x_1^{(0)}, \ldots, x_{n-1}^{(0)}),$$

$x_1^{(0)}, \ldots, x_n^{(0)}$ étant les coordonnées du point $\mathbf{x}_0$, et

$$F(\hat{\mathbf{x}}, \varphi(\hat{\mathbf{x}})) = 0$$

pour tous les points $\hat{\mathbf{x}} = (x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1}$ d'un voisinage $\hat{U}_0$ de $\hat{\mathbf{x}}_0 \in \mathbb{R}^{n-1}$. On dérive la dernière identité par rapport à $x_1, \ldots, x_{n-1}$ et on pose $\hat{\mathbf{x}} = \hat{\mathbf{x}}_0$, il vient en vertu de (11)

$$\frac{\partial \varphi}{\partial x_1}(\hat{\mathbf{x}}_0) = 0, \ldots, \frac{\partial \varphi}{\partial x_{n-1}}(\hat{x}_0) = 0.$$

Soit $\delta > 0$ un nombre positif suffisamment petit pour que le point $\hat{\mathbf{x}}_0 + \hat{\mathbf{a}}t$, avec $\hat{\mathbf{a}} = (a_1, \ldots, a_{n-1})$, soit pour $|t| \leqslant \delta$ dans $\hat{U}_0$. Dans ce cas, les formules

$$\hat{\mathbf{x}}(t) = \hat{\mathbf{x}}_0 + \hat{\mathbf{a}}t, \qquad x_n(t) = \varphi(\hat{\mathbf{x}}_0 + \hat{\mathbf{a}}t)$$

définissent dans $\mathscr{E}$ une courbe $t \mapsto \mathbf{x}(t)$, $|t| \leqslant \delta$, qui est sur l'hypersurface (7) et passe pour $t = 0$ par le point $\mathbf{x}_0$. De plus,

$$\hat{\mathbf{x}}'(0) = \hat{\mathbf{a}}$$

et

$$x_n'(0) = \left.\frac{d\varphi(\hat{\mathbf{x}}_0 + \hat{\mathbf{a}}t)}{dt}\right|_{t=0} = \frac{\partial \varphi}{\partial x_1}(\hat{\mathbf{x}}_0)\, a_1 + \ldots + \frac{\partial \varphi}{\partial x_{n-1}}(\hat{\mathbf{x}}_0)\, a_{n-1} = 0,$$

i.e. $x_n'(0) = a_n$. Par conséquent, $\mathbf{x}'(0) = \mathbf{a}$ et, partant, $\mathbf{a} \in \mathscr{H}_{\mathbf{x}_0}$. □

Définition 6. Un hyperplan de l'espace $\mathscr{E}$ parallèle au sous-espace $\mathscr{H}_{\mathbf{x}_0}$ par le point $\mathbf{x}_0$ s'appelle l'*hyperplan tangent* à l'hypersurface (7) au point $\mathbf{x}_0$.

D'après la proposition 1, l'équation de l'hyperplan tangent s'écrit

$$(\mathbf{x} - \mathbf{x}_0) \operatorname{grad} F(\mathbf{x}_0) = 0,$$

i.e.

$$\left(\frac{\partial F}{\partial x_1}\right)_0 (x - x_1) + \ldots + \left(\frac{\partial F}{\partial x_n}\right)_0 (x - x_n) = 0.$$

Le vecteur grad $F(\mathbf{x}_0)$ est orthogonal à l'hyperplan tangent.

Si $n = 2$, on retrouve un résultat connu d'Analyse, à savoir dans le plan, la tangente à une courbe quelconque

$$F(x, y) = 0$$

en son point régulier (x_0, y_0) a pour équation

$$\left(\frac{\partial F}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial F}{\partial y}\right)_0 (y - y_0) = 0.$$

* * *

On enseigne en Analyse que la *longueur* d'une courbe continue (1) est la limite (quand elle existe) des longueurs des segments de droite sous-tendant les arcs de courbe (l'espace $\mathscr{E}$ est supposé euclidien). Cette limite existe nécessairement pour toute courbe (1) différentiable (dont on dit qu'elle est *rectifiable*) et a la forme de l'intégrale

$$\int_a^b |\mathbf{x}'(t)|\, dt \tag{12}$$

En fait, on utilise la formule (12) sans donner la définition qui la précède (c'est au moins le cas des courbes différentiables). Aussi le procédé le plus simple consiste à accepter l'intégrale (12) en qualité de *définition* de la longueur d'une courbe différentiable et à ne s'adresser aux segments de droite que pour la légitimer heuristiquement. Ce procédé sera le nôtre pendant le troisième semestre quand nous nous placerons dans des cas analogues mais plus compliqués (quand nous aurons, par exemple, à définir l'aire d'une surface).

Soit

$$s(t) = \int_a^t |\mathbf{x}'(t)|\, dt \tag{13}$$

la longueur de la portion de courbe (1) comprise entre a et t. Si (1) est régulière, alors

$$s'(t) = |x'(t)| > 0,$$

ce qui permet le changement de paramètre $t \mapsto s(t)$. Ainsi, *toute courbe régulière est équivalente à une courbe paramétrée par la longueur*

de l'arc de courbe. On dit d'ordinaire de ces dernières courbes qu'elles *sont rapportées au paramètre naturel s.*

On supposera partout dans la suite que les courbes étudiées sont rapportées au paramètre naturel. La chose n'est pas foncièrement importante, mais les calculs se trouvent sensiblement simplifiés.

La dérivation par rapport à s sera symbolisée par un point :

$$\dot{\mathbf{x}}(s) = \frac{d\mathbf{x}(s)}{ds}, \quad \ddot{\mathbf{x}}(s) = \frac{d^2\mathbf{x}(s)}{ds^2}, \quad \text{et ainsi de suite.}$$

Si $t = s$, on a par la formule (13) :

$$\int_a^s |\dot{\mathbf{x}}(s)|\, ds = s$$

(et $a = 0$), d'où

$$|\dot{\mathbf{x}}(s)| = 1 \quad \text{pour tous les } s.$$

Inversement, si $|\dot{\mathbf{x}}(t)| = 1$ et $a = 0$, alors $t = s$.

Lemme 1. *Soit* $s \mapsto \mathbf{u}(s)$ *une fonction différentiable à valeurs vectorielles telle que* $|\mathbf{u}(s)| = 1$ *quel que soit s. On a*

$$\mathbf{u}(s)\,\dot{\mathbf{u}}(s) = 0 \quad \textit{pour tout } s. \tag{14}$$

Démonstration. Il suffit de remarquer que le produit scalaire (et le produit vectoriel d'ailleurs) de fonctions à valeurs vectorielles obéit à la règle usuelle de dérivation du produit de fonctions à valeurs numériques (parce que la démonstration classique reste entière pour ce cas). Aussi on dérive $\mathbf{u}(s)^2 = 1$ (on simplifie par 2) et on obtient (14). □

En particulier,

$$\dot{\mathbf{x}}(s)\,\ddot{\mathbf{x}}(s) = 0 \quad \text{quel que soit } s.$$

Cette formule importante nous sera utile plus d'une fois.

* * *

Plaçons-nous dans le cas particulier de courbes dans le plan. On notera comme toujours x, y les coordonnées (rectangulaires) dans le plan. Le rayon vecteur de coordonnées x, y sera désigné par $\mathbf{r}$ (au lieu du symbole $\mathbf{x}$ du cas général). Le symbole $\mathbf{t}(s)$ désignera pour toute courbe plane $\mathbf{r} = \mathbf{r}(s)$ (rapportée au paramètre naturel s) un vecteur tangent à la courbe au point $\mathbf{r}(s)$:

$$\mathbf{t}(s) = \dot{\mathbf{r}}(s).$$

On sait déjà que ce vecteur est unitaire et que

$$\mathbf{t}(s)\,\dot{\mathbf{t}}(s) = 0 \quad \text{pour tout } s.$$

Définition 7. On appelle *courbure* de la courbe $\mathbf{r} = \mathbf{r}(s)$ au point s, et on note $k(s)$, la longueur du vecteur $\dot{\mathbf{t}}(s)$.

Ainsi,

$$k(s) = |\dot{\mathbf{t}}(s)| = \sqrt{\ddot{x}^2(s) + \ddot{y}^2(s)}.$$

La courbure d'une courbe paramétrée par t quelconque est par définition la courbure d'une courbe équivalente rapportée au paramètre naturel. Sa formule obtenue par des calculs simples mais assez laborieux à l'aide des seules formules de dérivation de fonctions est suffisamment compliquée :

$$k = \left| \frac{x''y' - y''x'}{[(x')^2 + (y')^2]^{3/2}} \right|.$$

Intuitivement, le nombre $k(s)$ est la vitesse de rotation instantanée du vecteur unitaire $\mathbf{t}(s)$. Il est clair que cette vitesse est d'autant plus élevée que la courbe est recourbée, d'où le terme « courbure ».

On utilise des fois (dans le plan orienté) la *courbure relative* k_{rel} qui est égale à k si les vecteurs $\mathbf{t}$ et $\dot{\mathbf{t}}$ forment pour $k \neq 0$ un repère direct du plan et à $-k$ dans le cas contraire. Nous en aurons besoin dans la leçon 26.

Exemple 1. Si

$$x(s) = x_0 + sl, \quad y(s) = y_0 + sm, \quad \text{où } l^2 + m^2 = 1,$$

i.e. si la courbe considérée est une droite, alors $\ddot{x}(s) = 0$ et $\ddot{y}(s) = 0$. Aussi $k(s) = 0$ pour tous les s, i.e. *la courbure d'une droite est identiquement nulle*, comme il fallait s'y attendre. □

Vu que les fonctions linéaires sont, on le vérifie aisément, les seules fonctions à dérivée seconde identiquement nulle, la réciproque est également vraie, i.e. *une courbe à courbure identiquement nulle est une droite* (ou un segment de droite). □

On dit d'un point $\mathbf{r}_0 = \mathbf{r}(s_0)$ (tel que $k(s_0) = 0$) d'une courbe $\mathbf{r} = \mathbf{r}(s)$ que c'est un *point de courbure nulle*.

Exemple 2. Les équations paramétriques rapportées à s d'une circonférence de rayon R s'écrivent évidemment

$$x = R\cos\frac{s}{R}, \quad y = R\sin\frac{s}{R}.$$

Comme

$$\ddot{x} = -\frac{1}{R}\cos\frac{s}{R}, \quad \ddot{y} = -\frac{1}{R}\sin\frac{s}{R},$$

on a

$$k(s) = \frac{1}{R}.$$

Ainsi, *la courbure en tout point d'une circonférence est égale à l'inverse de son rayon.* □

La réciproque a également lieu, à savoir *une courbe à courbure constante est une circonférence* (ou un arc de cercle). □

Ce résultat découle d'un théorème général qui dit qu'*étant donnée une fonction quelconque* $k = k(s)$ (*définie et différentiable sur le segment* $|s| \leqslant s_0$), *si* s_0 *est suffisamment petit, il existe une courbe* $\mathbf{r} = \mathbf{r}(s)$, $|s| < s_0$, *à courbure* $k(s)$, *la courbe étant unique à une congruence près.* Nous nous passerons de la démonstration de ce résultat parce qu'un analogue en sera établi pour tout n dans la leçon suivante.

Si $k(s) \neq 0$, on définit le nombre $R(s) = \frac{1}{k(s)}$ appelé *rayon de courbure* de la courbe au point s.

Une courbe $\mathbf{r} = \mathbf{r}(s)$ telle que $k(s) \neq 0$ pour tout s est dite *de type général.* On définit en tous ses points le vecteur unitaire

$$\mathbf{n}(s) = \frac{\dot{\mathbf{t}}(s)}{k(s)}$$

dirigé suivant la *normale* à la courbe (i.e. suivant la droite perpendiculaire à la tangente au point de tangence).

Les vecteurs $\mathbf{t}(s)$ et $\mathbf{n}(s)$ forment pour tout s un repère orthonormé qu'on appelle le *repère mobile de Frenet* de la courbe.

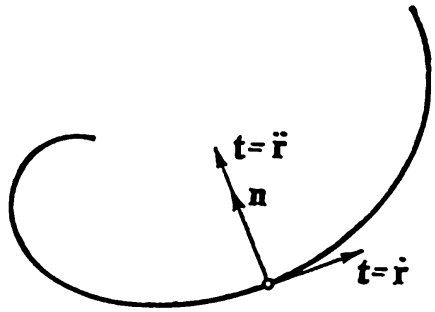
Par définition

$$\dot{\mathbf{t}}(s) = k(s)\,\mathbf{n}(s).$$

Cherchons son analogue pour $\dot{\mathbf{n}}(s)$. Soit

$$\dot{\mathbf{n}}(s) = \alpha(s)\,\mathbf{t}(s) + \beta(s)\,\mathbf{n}(s)$$

la décomposition de $\dot{\mathbf{n}}(s)$ dans le repère $\mathbf{t}(s)$, $\mathbf{n}(s)$. Comme $\mathbf{t}(s)\,\mathbf{n}(s) = 0$, on a $\dot{\mathbf{t}}(s)\,\mathbf{n}(s) + \mathbf{t}(s)\,\dot{\mathbf{n}}(s) = 0$, donc $\alpha(s) = \mathbf{t}(s)\,\dot{\mathbf{n}}(s) = -\dot{\mathbf{t}}(s)\,\mathbf{n}(s) = -k(s)$. D'autre part, $\beta(s) = \mathbf{n}(s)\,\dot{\mathbf{n}}(s) = 0$ par le



Repère de Frenet d'une courbe dans le plan

lemme 1. Ainsi, *toute courbe de type général vérifie les formules*

$$\dot{\mathbf{t}}(s) = k(s)\,\mathbf{n}(s), \qquad \dot{\mathbf{n}}(s) = -k(s)\,\mathbf{t}(s) \tag{15}$$

qui décrivent la rotation instantanée du repère mobile avec la variation de s. □

Les formules (15) constituent les *formules de Frenet pour une courbe plane.*

* * *

Plaçons-nous dans l'espace de dimension 3 et notons x, y, z les coordonnées dans l'espace et $\mathbf{r}$ le rayon vecteur des points. Le vecteur tangent $\dot{\mathbf{r}}(s)$ à toute courbe $\mathbf{r} = \mathbf{r}(s)$ (rapportée au paramètre naturel) sera toujours désigné par $\mathbf{t}(s)$. S'agissant des courbes dans l'espace, la longueur $|\dot{\mathbf{t}}(s)|$ du vecteur $\dot{\mathbf{t}}(s)$ s'appelle la *courbure* et se note $k(s)$ comme pour les courbes planes. Ainsi,

$$k(s) = \sqrt{\ddot{x}(s)^2 + \ddot{y}(s)^2 + \ddot{z}(s)^2}\,.$$

A l'instar du cas $n = 2$, une courbe $\mathbf{r} = \mathbf{r}(s)$ est dite *de type général* si $k(s) \neq 0$ quel que soit s. On définit alors le vecteur unitaire

$$\mathbf{n}(s) = \frac{\dot{\mathbf{t}}(s)}{k(s)}$$

appelé *vecteur de la normale principale* à la courbe.

Avec l'hypothèse d'espace orienté, on introduit un troisième vecteur $\mathbf{b}(s)$ tel que $\mathbf{b}(s)$, $\mathbf{t}(s)$ et $\mathbf{n}(s)$ constituent un repère orthonormé direct. Le vecteur $\mathbf{b}(s)$ est par définition le *vecteur de la binormale*, et le repère $\mathbf{t}(s)$, $\mathbf{n}(s)$, $\mathbf{b}(s)$ forme le *repère mobile de Frenet* de la courbe.

On a par construction

$$\dot{\mathbf{t}} = k\mathbf{n}$$

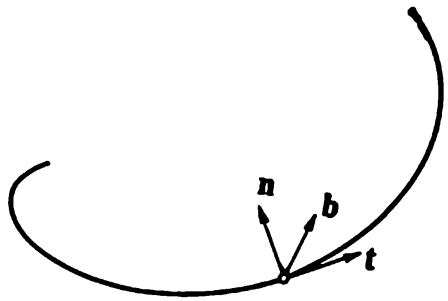
(on omet la variable s afin de simplifier les formules). Comme $\mathbf{b} = \mathbf{t} \times \mathbf{n}$, on a

$$\dot{\mathbf{b}} = \dot{\mathbf{t}} \times \mathbf{n} + \mathbf{t} \times \dot{\mathbf{n}} = \mathbf{t} \times \dot{\mathbf{n}},$$

d'où $\dot{\mathbf{b}}\mathbf{t} = 0$. Selon le lemme 1, $\dot{\mathbf{b}}\mathbf{b} = 0$, ce qui prouve la propriété « $\dot{\mathbf{b}}$ est colinéaire à $\mathbf{n}$ », i.e. il existe un nombre $\varkappa = \varkappa(s)$ pour lequel

$$\dot{\mathbf{b}} = -\varkappa\mathbf{n}.$$

Ce nombre porte le nom de *torsion* de la courbe au point s et représente la vitesse de rotation du vecteur de la binormale.



Repère de Frenet d'une courbe dans l'espace

Si l'on dérive les égalités $\mathbf{nt} = 0$ et $\mathbf{nb} = 0$, on trouve immédiatement $\dot{\mathbf{n}}\mathbf{t} = -\mathbf{n}\dot{\mathbf{t}} = -k$ et $\dot{\mathbf{n}}\mathbf{b} = -\mathbf{n}\dot{\mathbf{b}} = \varkappa$. Il est déjà connu que $\mathbf{n} = 0$ (lemme 1), donc

$$\dot{\mathbf{n}} = -k\mathbf{t} + \varkappa\mathbf{b}.$$

Ainsi, *toute courbe de type général vérifie les formules*

$$\begin{aligned} \dot{\mathbf{t}} &= k\mathbf{n}, \\ \dot{\mathbf{n}} &= -k\mathbf{t} + \varkappa\mathbf{b}, \\ \dot{\mathbf{b}} &= -\varkappa\mathbf{n}. \quad \square \end{aligned} \tag{16}$$

Ce sont *les formules de Frenet pour une courbe dans l'espace.*

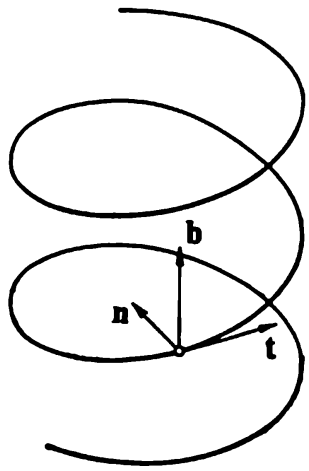
Exemple 1. Si la courbe $\mathbf{r} = \mathbf{r}(s)$ est dans le plan Π, alors les vecteurs $\dot{\mathbf{r}}(s)$ et $\ddot{\mathbf{r}}(s)$ sont parallèles à Π (du moment qu'il en est ainsi des accroissements $\mathbf{r}(s + \Delta s) - \mathbf{r}(s)$ et $\dot{\mathbf{r}}(s + \Delta s) - \dot{\mathbf{r}}(s)$ des vecteurs $\mathbf{r}(s)$ et $\dot{\mathbf{r}}(s)$). Aussi $\mathbf{t}(s)$, $\mathbf{n}(s) \parallel \Pi$ et, partant, $\mathbf{b}(s) \perp \Pi$. Cela prouve que $\mathbf{b}(s) = \text{const}$, donc $\varkappa(s) = 0$ pour tout s. Inversement, soit $\varkappa(s) = 0$ pour tout s, ce qui implique $\mathbf{b}(s) = \mathbf{b}_0 = \text{const}$. Dans ce cas, $(\mathbf{r}(s)\,\mathbf{b}_0)^{\cdot} = \mathbf{t}(s)\,\mathbf{b}_0 = 0$ quel que soit s, et $\mathbf{r}(s)\,\mathbf{b}_0$ est donc une constante. Cela signifie que la courbe $\mathbf{r} = \mathbf{r}(s)$ est dans le plan $\mathbf{r}\mathbf{b}_0 = \text{const}$. Ainsi, *une courbe dans l'espace est une courbe plane si et seulement si sa torsion est identiquement nulle.* $\square$

Exemple 2. Un point se déplaçant à une vitesse constante suivant la génératrice d'un cylindre circulaire droit animé d'un mouvement uniforme autour de son axe décrit une courbe appelée *hélice.* Les équations de l'hélice s'écrivent

$$x = a\cos t, \quad y = a\sin t, \quad z = bt.$$

On a

$$x' = -a\sin t, \quad y' = a\cos t, \quad z = b,$$



Hélice

d'où

$$s' = \sqrt{(x')^2 + (y')^2 + (z')^2} = \sqrt{a^2 + b^2}.$$

Ainsi, $s = ct$, avec $c = \sqrt{a^2 + b^2}$, donc

$$x = a \cos \frac{s}{c}, \quad y = a \sin \frac{s}{c}, \quad z = \frac{b}{c} s,$$

Comme

$$\dot{x} = -\frac{a}{c} \sin \frac{s}{c}, \quad \dot{y} = \frac{a}{c} \cos \frac{s}{c}, \quad \dot{z} = \frac{b}{c},$$

$$\ddot{x} = -\frac{a}{c^2} \cos \frac{s}{c}, \quad \ddot{y} = -\frac{a}{c^2} \sin \frac{s}{c}, \quad \ddot{z} = 0,$$

on a

$$k = \sqrt{\ddot{x}^2 + \ddot{y}^2 + \ddot{z}^2} = \frac{a}{c^2} = \text{const}$$

et

$$\mathbf{t} = -\frac{a}{c} \sin \frac{s}{c}\, \mathbf{i} + \frac{a}{c} \cos \frac{s}{c}\, \mathbf{j} + \frac{b}{c}\, \mathbf{k},$$

$$\mathbf{n} = -\cos \frac{s}{c}\, \mathbf{i} - \sin \frac{s}{c}\, \mathbf{j},$$

$$\mathbf{b} = \mathbf{t} \times \mathbf{n} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\frac{a}{c} \sin \frac{s}{c} & \frac{a}{c} \cos \frac{s}{c} & \frac{b}{c} \\ -\cos \frac{s}{c} & -\sin \frac{s}{c} & 0 \end{vmatrix} =$$

$$= \frac{b}{c} \sin \frac{s}{c}\, \mathbf{i} - \frac{b}{c} \cos \frac{s}{c}\, \mathbf{j} + \frac{a}{c}\, \mathbf{k}.$$

Aussi

$$\dot{\mathbf{b}} = \frac{b}{c^2} \cos \frac{s}{c}\, \mathbf{i} + \frac{b}{c^2} \sin \frac{s}{c}\, \mathbf{j} = -\frac{b}{c^2}\, \mathbf{n},$$

si bien que

$$\varkappa = \frac{b}{c^2} = \text{const.}$$

Ainsi, *la courbure et la torsion de l'hélice sont des constantes.* □

Selon un théorème général que nous démontrerons dans la leçon 24, *toute courbe à courbure constante et à torsion constante est une hélice* (*ou un arc d'hélice*), et vice-versa.

LEÇON 24

Projections d'une courbe sur les plans de coordonnées du repère mobile. — Formules de Frenet pour une courbe dans l'espace n-dimensionnel. — Une courbe définie par ses courbures. — Surfaces régulières. — Exemples de surfaces.

On se propose d'étudier le comportement d'une courbe gauche quelconque $\mathbf{r} = \mathbf{r}(s)$ au voisinage d'un de ses points. A cet effet, on prend ce point pour origine O, les vecteurs $\mathbf{t}_0$, $\mathbf{n}_0$, $\mathbf{b}_0$ du repère mobile au point O pour vecteurs de base $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, et on paramètre la courbe par s à partir de O. Dans ce cas,

$$\mathbf{r}(0) = 0, \quad \dot{\mathbf{r}}(0) = \mathbf{t}_0 = \mathbf{i}, \quad \ddot{\mathbf{r}}(0) = k_0\mathbf{n}_0 = k_0\mathbf{j},$$

$$\dddot{\mathbf{r}}(0) = (\dot{k})_0\,\mathbf{n}_0 + k_0\dot{\mathbf{n}}_0 = -k_0^2\mathbf{i} + (\dot{k})_0\,\mathbf{j} + k_0\varkappa_0\mathbf{k},$$

k_0, $(\dot{k})_0$ et $\varkappa_0$ étant les valeurs des fonctions k, $\dot{k}$ et $\varkappa$ pour $s = 0$. On a donc par la formule de Taylor

$$\mathbf{r}(s) = \mathbf{r}(0) + s\dot{\mathbf{r}}(0) + \frac{s^2}{2}\ddot{\mathbf{r}}(0) + \frac{s^3}{6}\dddot{\mathbf{r}}(0) + \ldots =$$

$$= \left(s - \frac{k_0^2}{6}s^3 + \ldots\right)\mathbf{i} + \left(\frac{k_0}{2}s^2 + \frac{(\dot{k})_0}{6}s^3 + \ldots\right)\mathbf{j} +$$

$$+ \left(\frac{k_0\varkappa_0}{6}s^3 + \ldots\right)\mathbf{k}.$$

Autrement dit, la courbe est définie au voisinage de O par les équations paramétriques

$$x = s + \ldots,$$

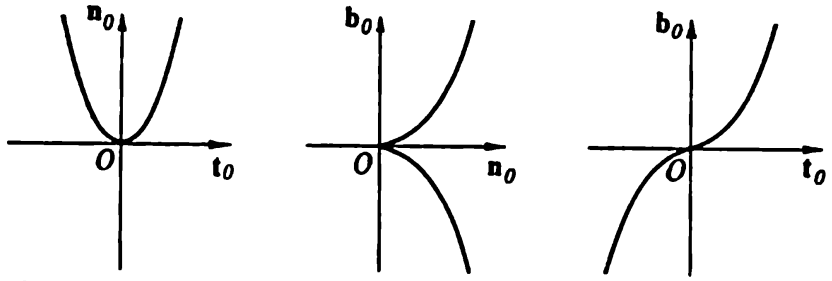
$$y = \frac{k_0}{2}s^2 + \ldots,$$

$$z = \frac{k_0\varkappa_0}{6}s^3 + \ldots$$

Si $k_0 \neq 0$, $\varkappa_0 \neq 0$, la projection de la courbe sur le plan $O\mathbf{ij} = O\mathbf{t}_0\mathbf{n}_0$ (appelé *plan osculateur* en O à la courbe) est approximativement la parabole

$$x = s, \quad y = \frac{k_0}{2}s^2,$$

sa projection sur le plan $O\mathbf{jk} = O\mathbf{n}_0\mathbf{b}_0$ (appelé *plan normal* à la courbe



Projection sur le plan osculateur *Projection sur le plan normal* *Projection sur le plan rectifiant*

en O) la parabole semi-cubique

$$y = \frac{k_0}{2} s^2, \quad z = \frac{k_0 \varkappa_0}{6} s^3$$

et sa projection sur le plan $O\mathbf{ik} = O\mathbf{t}_0\mathbf{b}_0$ (appelé *plan rectifiant* de la courbe en O) la parabole cubique

$$x = s, \quad z = \frac{k_0 \varkappa_0}{2} s^3.$$

Cela donne une idée assez nette de la forme d'une courbe gauche au voisinage de son point quelconque (où sa courbure et sa torsion sont non nulles).

* * *

Généralisons les résultats de la leçon précédente à n quelconque. Soit $\mathbf{x} = \mathbf{x}(s)$, $|s| \leqslant s_0$, une courbe quelconque (rapportée au paramètre naturel) dans l'espace euclidien orienté $\mathscr{E}$ de dimension n. On suppose que *les vecteurs*

$$\dot{\mathbf{x}}(s), \ldots, \overset{(n-1)}{\mathbf{x}}(s)$$

sont linéairement indépendants pour tout s (on dit de ces courbes qu'elles sont *de type général*) et on applique le procédé d'orthonormalisation de Gram-Schmidt, ce qui donne la famille orthonormée de vecteurs $\mathbf{t}_1(s), \ldots, \mathbf{t}_{n-1}(s)$. Soit $\mathbf{t}_n(s)$ le vecteur (bien défini) qui complète la famille de façon à former le repère orthonormé direct

$$(1) \qquad \mathbf{t}_1(s), \ldots, \mathbf{t}_{n-1}(s), \quad \mathbf{t}_n(s).$$

Définition 1. Le repère (1) s'appelle le *repère mobile de Frenet* de la courbe $\mathbf{x} = \mathbf{x}(s)$ de type général au point s.

Soit

$$\dot{\mathbf{t}}_i = \sum_{j=1}^{n} \alpha_{ij} \mathbf{t}_j, \quad i = 1, \ldots, n$$

(on omet la variable s afin de simplifier les formules). Par construction, le vecteur $\mathbf{t}_i$, $i = 1, \ldots, n-1$, s'exprime linéairement en fonction des vecteurs $\dot{\mathbf{x}}, \ldots, \overset{(i)}{\mathbf{x}}$, si bien que $\dot{\mathbf{t}}_i$ est combinaison linéaire des $\dot{\mathbf{x}}, \ldots, \overset{(i+1)}{\mathbf{x}}$. Comme $\dot{\mathbf{x}}, \ldots, \overset{(i+1)}{\mathbf{x}}$ dépendent linéairement de $\mathbf{t}_1, \ldots, \mathbf{t}_{i+1}$, cela prouve $\alpha_{ij} = 0$ pour $j > i + 1$.

D'autre part, $\mathbf{t}_i\mathbf{t}_j = \delta_{ij}$ entraîne $\dot{\mathbf{t}}_i\mathbf{t}_j + \mathbf{t}_i\dot{\mathbf{t}}_j = 0$, i.e.

$$\alpha_{ij} + \alpha_{ji} = 0.$$

Aussi $\alpha_{ii} = 0$ et $\alpha_{ij} = 0$ pour $j < i - 1$.

Ainsi, seuls les coefficients $\alpha_{i,\,i+1} = -\alpha_{i+1,\,i}$ sont différents de 0. On pose

$$k_1 = \alpha_{12}, \quad k_2 = \alpha_{23}, \ \ldots, k_{n-1} = \alpha_{n-1,n},$$

et on constate la validité des formules

$$(2) \qquad \begin{aligned} \dot{\mathbf{t}}_1 &= k_1\mathbf{t}_2, \\ \dot{\mathbf{t}}_2 &= -k_1\mathbf{t}_1 + k_2\mathbf{t}_3, \\ &\ldots\ldots\ldots\ldots \\ \dot{\mathbf{t}}_{n-1} &= -k_{n-2}\mathbf{t}_{n-2} + k_{n-1}\mathbf{t}_n, \\ \dot{\mathbf{t}}_n &= -k_{n-1}\mathbf{t}_{n-1}. \end{aligned}$$

Ce sont *les formules de Frenet pour une courbe dans l'espace de dimension n.*

Les fonctions $k_1 = k_1(s), \ldots, k_{n-1} = k_{n-1}(s)$ s'appellent les *courbures* de la courbe. On note qu'elles sont définies seulement pour une courbe de type général.

Dans les formules

$$(3) \qquad \mathbf{t}_i = \beta_{i1}\dot{\mathbf{x}} + \ldots + \beta_{ii}\overset{(i)}{\mathbf{x}}, \quad i = 1, \ldots, n-1,$$

obtenues au cours de l'orthonormalisation de Gram-Schmidt, les derniers coefficients β_{ii} sont positifs. Aussi il en est également des $\gamma_{ii} = \beta_{ii}^{-1}$ des formules inverses

$$(4) \qquad \overset{(i)}{\mathbf{x}} = \gamma_{i1}\mathbf{t}_1 + \ldots + \gamma_{ii}\mathbf{t}_i.$$

On dérive (3), il vient

$$\dot{\mathbf{t}}_i = \dot{\beta}_{i1}\dot{\mathbf{x}} + (\dot{\beta}_{i2} + \beta_{i1})\ddot{\mathbf{x}} + \ldots + (\dot{\beta}_{ii} + \beta_{i,i-1})\overset{(i)}{\mathbf{x}} + \beta_{ii}\overset{(i+1)}{\mathbf{x}},$$
$$i = 1, \ldots, n-1.$$

Si l'on substitue pour $i < n - 1$ aux vecteurs $\dot{\mathbf{x}}, \ldots, \overset{(i+1)}{\mathbf{x}}$ leurs expressions (4), on a les formules (2), ce qui montre que

$$k_i = \beta_{ii}\gamma_{i+1,\,i+1} \quad \text{pour tout} \quad i = 1, \ldots, n-2.$$

Ainsi, *les courbures*

$$k_1, \ldots, k_{n-2}$$

de toute courbe de type général sont positives. Quant à k_{n-1} (un analogue de la torsion), elle peut être de signe quelconque.

* * *

On montre que n'importe quelles $n - 1$ fonctions

$$k_1(s) > 0, \ldots, k_{n-2}(s) > 0, \; k_{n-1}(s) \tag{5}$$

sont les courbures d'une courbe et qu'elles la définissent de façon unique (à une congruence près).

Théorème 1. *Soient, sur un segment* $|s| \leqslant s_0$ *quelconque,* $n - 1$ *fonctions différentiables* (5) *qui sont toutes positives, sauf peut-être la dernière. Pour toute origine* $O \in \mathscr{E}$ *et tout repère orthonormé direct* $\mathbf{i}_1, \ldots, \mathbf{i}_n$, *il existe une, et une seule, courbe régulière* $\mathbf{x} = \mathbf{x}(s)$, $|s| \leqslant s_0$, *de type général qui jouit de deux propriétés suivantes:*

1) *ses courbures sont les fonctions données* (5);

2) *on a pour* $s = 0$

$$\mathbf{x}(0) = \mathbf{0}, \quad \mathbf{t}_1(0) = \mathbf{i}_1, \ldots, \mathbf{t}_n(0) = \mathbf{i}_n.$$

Démonstration. On procédera par étapes.

Etape première. Nous allons utiliser le résultat général connu sous le nom de *théorème d'existence et d'unicité des solutions des équations différentielles linéaires* (dont la démonstration sera faite pendant le troisième semestre dans le cadre de la théorie des équations différentielles).

Théorème d'existence et d'unicité. *Soit, sur un segment quelconque* $|s| \leqslant s_0$, m^2 *fonctions différentiables* $A_{ij}(s)$, $i, j = 1, \ldots, m$, *et soit* $x_1^{(0)}, \ldots, x_m^{(0)}$ *des nombres arbitraires. Il existe une, et une seule, famille de fonctions différentiables* $x_1(s), \ldots, x_m(s)$, $|s| \leqslant s_0$, *telles que*

1) *on ait identiquement en* s, $|s| \leqslant s_0$,

$$\begin{aligned} \dot{x}_1 &= A_{11}x_1 + \ldots + A_{1m}x_m, \\ &\ldots\ldots\ldots\ldots\ldots \\ \dot{x}_m &= A_{m1}x_1 + \ldots + A_{mm}x_m; \end{aligned} \tag{6}$$

2) *on ait pour* $s = 0$

$$x_1(0) = x_1^{(0)}, \ldots, x_m(0) = x_m^{(0)}. \quad \square$$

Dans le cas de relations (2) qui sont pour $k_1, \ldots, k_{n-1}$ donnés des équations de la forme (6) pour $m = n^2$ coordonnées des vecteurs $\mathbf{t}_1, \ldots, \mathbf{t}_n$, le segment $|s| \leqslant s_0$ contient par ce théorème une famille unique de fonctions à valeurs vectorielles $\mathbf{t}_1 = \mathbf{t}_1(s), \ldots, \mathbf{t}_n = \mathbf{t}_{n-1}(s)$ telles que

1) on ait pour tout s les relations (2);

2) on ait pour $s = 0$

$$\mathbf{t}_1(0) = \mathbf{i}_1, \ldots, \mathbf{t}_n(0) = \mathbf{i}_n. \tag{7}$$

Etape 2. Considérons les produits scalaires $\mathbf{t}_i\mathbf{t}_j$, $i, j = 1, \ldots, n$. D'après les relations (2)

$$(\mathbf{t}_i\mathbf{t}_j)^{\cdot} = \dot{\mathbf{t}}_i\mathbf{t}_j + \dot{\mathbf{t}}_i\mathbf{t}_j = (-k_{i-1}\mathbf{t}_{i-1} + k_i\mathbf{t}_{i+1})\,\mathbf{t}_j + \mathbf{t}_i\,(-k_{j-1}\mathbf{t}_{j-1} + k_j\mathbf{t}_{j+1})$$

(on pose par convention $\mathbf{t}_0 = 0$ et $\mathbf{t}_{n+1} = 0$), i.e.

$$(\mathbf{t}_i\mathbf{t}_j)^{\cdot} = -k_{i-1}(\mathbf{t}_{i-1}\mathbf{t}_j) + k_i(\mathbf{t}_{i+1}\mathbf{t}_j) - k_{j-1}(\mathbf{t}_i\mathbf{t}_{j-1}) + k_j(\mathbf{t}_i\mathbf{t}_{j+1}). \tag{8}$$

égalités qu'on peut considérer comme équations (6) pour $m = \frac{n(n+1)}{2}$ fonctions $\mathbf{t}_i\mathbf{t}_j$. Aussi il existe, selon le théorème d'existence et d'unicité, un ensemble unique de ces fonctions ayant la propriété d'être égales à $\delta_{ij} = \mathbf{i}_i\mathbf{i}_j$ pour $s = 0$ (i.e. elles sont nulles si $i \neq j$ et valent 1 dans le cas contraire).

D'autre part, il est immédiat de vérifier que les équations (8) sont satisfaites par $\mathbf{t}_i\mathbf{t}_j \equiv \delta_{ij}$. (En effet, si $i \neq j-1, j+1$, la somme $-k_{i-1}\delta_{i-1,\,j} + k_i\delta_{i+1,\,j} - k_{j-1}\delta_{i,\,j-1} + k_j\delta_{i,\,j+1}$ a tous ses termes nuls, et si $i = j-1, j+1$, cette somme comprend seulement deux termes non nuls qui se réduisent mutuellement.) Par notre théorème, on a donc pour tout s les égalités $\mathbf{t}_i\mathbf{t}_j = \delta_{ij}$, $i, j = 1, \ldots, n$, qui signifient que les vecteurs $\mathbf{t}_1, \ldots, \mathbf{t}_n$ constituent pour tout s, $|s| \leqslant s_0$, un repère orthonormé.

Ce repère coïncidant pour $s = 0$ avec $\mathbf{i}_1, \ldots, \mathbf{i}_n$ direct est direct quel que soit s.

Etape 3. Soit les dérivées successives du vecteur $\mathbf{t}_1$:

$$\mathbf{t}_1,\ \dot{\mathbf{t}}_1,\ \ddot{\mathbf{t}}_1,\ \ldots,\ \overset{(n-1)}{\mathbf{t}_1}. \tag{9}$$

On emploie le procédé d'orthonormalisation de Gram-Schmidt. Comme $\mathbf{t}_1$ est un vecteur unitaire, on ne fait rien au premier pas. Le vecteur $\dot{\mathbf{t}}_1$ étant orthogonal à $\mathbf{t}_1$ (par le lemme 1 de la leçon précédente), le deuxième pas consiste à le normer. Le vecteur $\mathbf{t}_2$ est unitaire par les résultats obtenus et $k_1 > 0$ par hypothèse, si bien que $|\dot{\mathbf{t}}_1| = k_1$ par suite de la première relation (2). On obtient donc au deuxième pas le vecteur

$$\mathbf{t}_2 = \frac{\dot{\mathbf{t}}_1}{k_1}.$$

Au troisième pas, on considère le vecteur

$$\ddot{\mathbf{t}}_1 = (k_1\mathbf{t}_2)^{\cdot} = \dot{k}_1\mathbf{t}_2 + k_1\dot{\mathbf{t}}_2 = -k_1^2\mathbf{t}_1 + \dot{k}_1\mathbf{t}_2 + k_1k_2\mathbf{t}_3,$$

on en retranche une combinaison linéaire de $\mathbf{t}_1$ et $\mathbf{t}_2$ telle que le résultat soit un vecteur orthogonal à ceux-ci, puis on norme ce vecteur orthogonal. Il est déjà connu que $\mathbf{t}_1$, $\mathbf{t}_2$, $\mathbf{t}_3$ forment une famille orthonormée et $k_1k_2 > 0$ par hypothèse, si bien que le processus aboutit évidemment à $\mathbf{t}_3$.

Il est clair que ces raisonnements sont de caractère général et que chaque pas de l'orthonormalisation donne donc $\mathbf{t}_i$, $i = 1, \ldots, n-1$, correspondant. Ainsi, la famille $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_{n-1}$ est caractérisée de façon unique comme famille orthonormée de vecteurs obtenue de (9) par le procédé de Gram-Schmidt.

Etape 4. Soit

$$\mathbf{x}(s) = \int_0^s \mathbf{t}_1(s)\,ds, \qquad |s| \leqslant s_0, \tag{10}$$

auquel cas $\mathbf{x}(0) = 0$ et $\dot{\mathbf{x}}(s) = \mathbf{t}_1(s)$, i.e. la courbe $\mathbf{x} = \mathbf{x}(s)$, $|s| \leqslant s_0$, a O pour point initial et admet en $\mathbf{x}(s)$ un vecteur tangent $\mathbf{t}_1(s)$. Quelle que soit la courbe, les $n-1$ premiers vecteurs du repère mobile proviennent par Gram-Schmidt des $n-1$ premières dérivées du vecteur tangent. Les résultats ci-dessus autorisent donc à les identifier avec les vecteurs $\mathbf{t}_1, \ldots, \mathbf{t}_{n-1}$.

Quant au dernier vecteur du repère mobile, il est bien caractérisé comme étant un vecteur unitaire qui forme ensemble avec les $n-1$ premiers vecteurs un repère direct. Le repère $\mathbf{t}_1, \ldots, \mathbf{t}_{n-1}, \mathbf{t}_n$ étant direct, ce doit être $\mathbf{t}_n$.

Ainsi, on vient d'établir que les vecteurs $\mathbf{t}_1(s), \ldots, \mathbf{t}_n(s)$ forment pour tout s le repère mobile de la courbe $\mathbf{x} = \mathbf{x}(s)$. Comme ces vecteurs vérifient les formules de Frenet (2), les fonctions $k_i(s)$, $i = 1, \ldots, n-1$, y figurant sont les courbures de $\mathbf{x} = \mathbf{x}(s)$.

Cela démontre complètement l'existence d'une courbe $\mathbf{x} = \mathbf{x}(s)$ jouissant des propriétés 1) et 2).

La courbe est unique. En effet, selon le théorème d'existence et d'unicité, le repère mobile $\mathbf{t}_1(s), \ldots, \mathbf{t}_n(s)$ est bien défini par les équations (2) et les conditions initiales (7), et le rayon vecteur $\mathbf{x}(s)$ l'est (par la formule (10)) moyennant la relation $\dot{\mathbf{x}}(s) = \mathbf{t}_1(s)$ et la condition initiale $\mathbf{x}(0) = 0$. □

* * *

Par analogie avec les définitions 1 à 3 de la leçon 23, on énonce pour tout k, $0 < k < n$, la définition « paramétrique » d'une surface de dimension k dans l'espace n-dimensionnel. Le souci de la simplicité nous fait nous borner à $k = 2$ et à $n = 3$.

Soit W un ouvert quelconque de l'espace $\mathbb{R}^2$ de dimension 2, dont les points sont des couples (u, v) de nombres réels. Une application $W \to \mathscr{E}$ quelconque de W dans l'espace euclidien $\mathscr{E}$ de dimension 3 est définie (si $\mathscr{E}$ est muni d'une origine O), ou bien par la donnée d'une fonction à valeurs vectorielles $\mathbf{r} = \mathbf{r}(u, v)$ de domaine de définition W, ou bien par la donnée (si l'on fixe dans $\mathscr{E}$ les coordonnées rectangulaires x, y, z) de trois fonctions numériques

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v). \tag{11}$$

On utilisera une fois de plus les seules applications $(u, v) \mapsto \mathbf{r}(u, v)$ *différentiables*, i.e. des applications telles que les fonctions (11) soient différentiables dans W. Cela définit donc les dérivées partielles

$$\mathbf{r}_u = x_u\mathbf{i} + y_u\mathbf{j} + z_u\mathbf{k}, \quad \mathbf{r}_v = x_v\mathbf{i} + y_v\mathbf{i} + z_v\mathbf{k} \tag{12}$$

(on omet les variables u, v afin de simplifier l'écriture).

Définition 2. Une application $(u, v) \mapsto \mathbf{r}(u, v)$ s'appelle une *surface régulière* si les vecteurs (12) sont linéairement indépendants pour tout point $(u, v) \in W$.

L'ensemble des points de $\mathscr{E}$ de rayons vecteurs $\mathbf{r}(u, v)$, $(u, v) \in \in W$, constitue le *support* de la surface $\mathbf{r} = \mathbf{r}(u, v)$.

On rappelle qu'en Analyse on appelle *difféomorphisme* une application bijective $W \to W_1$ d'un ouvert $W \subset \mathbb{R}^2$ sur un ouvert $W_1 \subset \subset \mathbb{R}^2$ si les fonctions

$$u_1 = u_1(u, v), \quad v_1 = v_1(u, v) \tag{13}$$

et

$$u = u(u_1, v_1), \quad v = v(u_1, v_1) \tag{14}$$

qui définissent $W \to W_1$ et l'application réciproque $W_1 \to W$ sont différentiables. Une bijection définie par (13) différentiables est un difféomorphisme si et seulement si son *jacobien*

$$\frac{\partial(u_1, v_1)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial u_1}{\partial u} & \dfrac{\partial u_1}{\partial v} \\ \dfrac{\partial v_1}{\partial u} & \dfrac{\partial v_1}{\partial v} \end{vmatrix} \tag{15}$$

est non nul partout dans W. Si le jacobien (15) des fonctions (13) (qu'on ne suppose pas à priori définissant une bijection) est non nul partout dans W, ces fonctions définissent un *difféomorphisme local*, i.e. il existe un voisinage $U \subset W$ de tout point $(u_0, v_0) \in W$ sur lequel cette application est un difféomorphisme de U sur un voisinage $U_1 \subset W_1$ du point $(u_1(u_0, v_0), v_1(u_0, v_0))$ (*théorème d'application réciproque*).

Soient deux surfaces

$$\mathbf{r} = \mathbf{r}(u, v), \quad (u, v) \in W, \tag{16}$$

et

$$\mathbf{r} = \mathbf{r}_1(u_1, v_1), \quad (u_1, v_1) \in W_1. \tag{17}$$

Définition 3. Les surfaces (16) et (17) sont dites *équivalentes* s'il existe un difféomorphisme

$$u_1 = u_1(u, v), \quad v_1 = v_1(u, v)$$

de l'ouvert W sur l'ouvert W_1 tel qu'on ait

$$\mathbf{r}(u, v) = \mathbf{r}_1(u_1(u, v), v_1(u, v))$$

pour tout point $(u, v) \in W$.

Il est clair que deux surfaces équivalentes ont même support.

Toute fonction différentiable $z = z(x, y)$ de deux variables définie dans W détermine par la formule

$$\mathbf{r}(u, v) = u\mathbf{i} + v\mathbf{j} + z(u, v)\mathbf{k}$$

une surface régulière qui en est le *graphe*.

Il se trouve qu'avec les axes de coordonnées choisis convenablement, *toute surface régulière est localement équivalente au graphe d'une fonction différentiable*. En effet, l'indépendance linéaire des vecteurs $\mathbf{r}_u$, $\mathbf{r}_v$ entraîne que la matrice

$$\begin{pmatrix} x_u & y_u & z_u \\ x_v & y_v & z_v \end{pmatrix}$$

est de rang 2 en tout point $(u_0, v_0) \in W$, i.e. au moins l'un de ses mineurs d'ordre 2 est non nul. Soit, pour fixer les idées,

$$\begin{vmatrix} x_u & y_u \\ x_v & y_v \end{vmatrix} \neq 0.$$

D'après le théorème d'application réciproque (relatif aux fonctions $x = x(u, v)$ et $y = y(u, v)$), il existe un voisinage $U_0 \subset W$ du point (u_0, v_0) et un voisinage $V_0 \subset \mathbb{R}^2$ du point $(x_0, y_0) \in \mathbb{R}^2$, où $x_0 = x(u_0, v_0)$, $y_0 = y(u_0, v_0)$, tels que les fonctions $x = x(u, v)$ et $y = y(u, v)$ définissent un difféomorphisme de U_0 sur V_0. Si

$$u = u(x, y), \quad v = v(x, y)$$

définissent le difféomorphisme réciproque, alors la surface $\mathbf{r} = \mathbf{r}(u, v)$ est équivalente sur U_0 au graphe de la fonction $z = z(u(x, y), v(x, y))$. □

Bien qu'une surface ne constitue pas un ensemble, on l'identifie souvent en paroles avec son support. On dit par exemple des points d'un support que ce sont les *points de la surface*, et ainsi de suite.

Une surface régulière peut en général être une application non injective dans $\mathscr{E}$ (elle présente des points, des courbes et des aires « en chevauchement »), mais seules nous intéresseront pendant ce semestre des portions de surface qui soient suffisamment petites pour que la surface ainsi limitée soit équivalente à un graphe et constitue donc une *application injective*.

Si un point M de la surface a le rayon vecteur $\mathbf{r}(u, v)$, alors les nombres u et v sont les *coordonnées de M sur la surface*. La définition est correcte par suite de la propriété de l'application $(u, v) \mapsto \mathbf{r}(u, v)$ d'être injective.

Toute courbe

$$u = u(t), \quad v = v(t) \tag{18}$$

du domaine W définit une courbe

$$\mathbf{r} = \mathbf{r}(u(t), v(t)) \tag{19}$$

de l'espace $\mathscr{E}$ dont on dit qu'elle est *sur la surface* $\mathbf{r} = \mathbf{r}(u, v)$. Les équations (18) s'appellent les équations de la courbe (19) rapportée aux coordonnées u, v sur la surface.

En particulier, on définit sur une surface les courbes $u = \text{const}$ et $v = \text{const}$ appelées les *lignes de coordonnées*. Ces courbes forment un *réseau de coordonnées*.

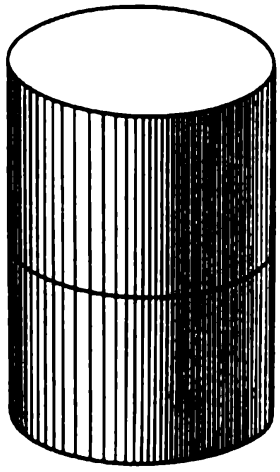
* * *

Exemples de surfaces.

1. Le support de la surface

$$x = R \cos u, \quad y = R \sin u, \quad z = v \tag{20}$$

est un cylindre circulaire droit de rayon R, ce qui explique le nom de *cylindre* (circulaire) donné à la surface (20) même.



Cylindre circulaire

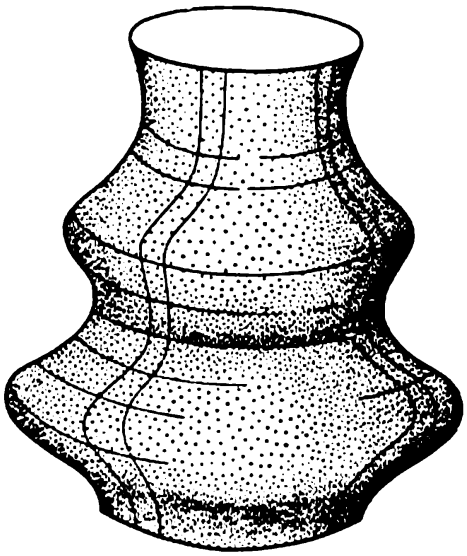
Quand $-\infty < u < +\infty$, chaque point du cylindre est recouvert un nombre infini (dénombrable) de fois par les points du plan (u, v). Si l'on veut que l'application soit injective, il faut admettre que $0 < u < 2\pi$, ce qui correspond au cylindre pourvu « d'une fente». Du moment que nous nous intéressons à l'aspect local du problème, nous négligerons les phénomènes de cette nature.

Le réseau de coordonnées du cylindre comprend les droites « verticales» $u = \text{const}$ et les circonférences « horizontales» $v = \text{const}$.

2. Soit $x = x(v)$, $z = z(v)$ une courbe régulière quelconque dans le plan Oxz, qui ne coupe pas l'axe Oz. La surface

$$x = x(v)\cos u, \quad y = x(v)\sin u, \quad z = z(v) \tag{21}$$

s'appelle *surface de révolution*, et la courbe $x = x(v)$, $z = z(v)$ en



Surface de révolution

constitue le *profil*. Intuitivement, la surface (21) est décrite par son profil qui tourne autour de l'axe Oz.

La régularité de (21), i.e. l'indépendance linéaire des vecteurs

$$\mathbf{r}_u = (-x(v)\sin u,\ x(v)\cos u,\ 0),$$
$$\mathbf{r}_v = (x'(v)\cos u,\ x'(v)\sin u,\ z'(v)),$$

est garantie par la régularité du profil (i.e. par la condition $x'(v)^2 + z'(v)^2 \neq 0$) et par $x(v) \neq 0$ (ce qui signifie que le profil ne coupe pas l'axe de rotation Oz).

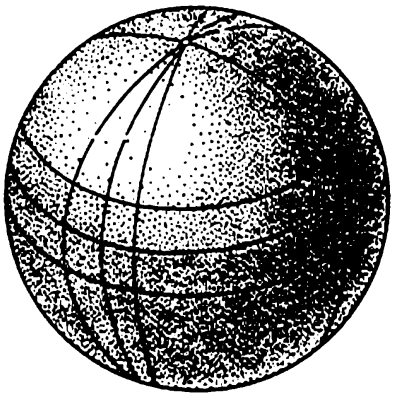
Le réseau de coordonnées sur la surface (21) est formé de courbes engendrées par le profil tournant autour de Oz (on les appelle les *méridiens*) et de circonférences qui leur sont perpendiculaires (ce sont les *parallèles*).

Le cylindre est une surface de révolution dont le profil est la droite $x = R$, $z = v$.

La surface de révolution de profil $x = R \cos v$, $z = R \sin v$ (une circonférence) est la *sphère*

$$x = R \cos v \cos u, \quad y = R \cos v \sin u, \quad z = R \sin v$$

de rayon R et de centre en le point O. Les coordonnées u et v sont ici les « coordonnées géographiques» connues (la longitude et la latitude), et les lignes de coordonnées représentent les méridiens et les parallèles des cartes géographiques.



Sphère

On note qu'on devrait, à vrai dire, se limiter à une portion de la circonférence $x = R \cos v$, $z = R \sin v$, qui ne coupe pas l'axe Oz et, partant, à la sphère privée de ses pôles. En effet, les coordonnées u et v n'ont aucun sens aux pôles. Mais d'après nos conventions, de tels faits ne sont pas pris en considération.

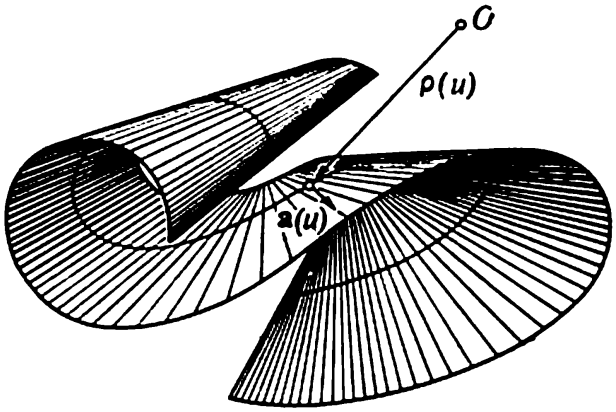
3. Une surface $\mathbf{r} = \mathbf{r}(u, v)$ est dite *réglée* si

$$\mathbf{r}(u, v) = \boldsymbol{\rho}(u) + v\mathbf{a}(u), \tag{22}$$

$\boldsymbol{\rho}(u)$ et $\mathbf{a}(u)$ étant deux fonctions à valeurs vectorielles quelconques telles que $\boldsymbol{\rho}'(u) + v\mathbf{a}'(u)$ et $\mathbf{a}(u)$ soient pour tous les u et v considérés linéairement indépendants, si bien qu'on a, en particulier, $\mathbf{a}(u) \neq 0$ quel que soit u (cette propriété des fonctions $\boldsymbol{\rho}(u)$ et $\mathbf{a}(u)$ garantit la régularité). La courbe $u = u_0 = \text{const}$ est une droite de vecteur directeur $\mathbf{a}(u_0)$ qui passe par un point de rayon vecteur $\boldsymbol{\rho}(u_0)$. Ainsi, on se représente intuitivement une surface réglée comme une surface engendrée par une droite qui se déplace dans l'espace. Cf. I.23, déf. 1.

Il est clair qu'on ne nuit nullement à la généralité de l'exposé si l'on prend $\mathbf{a}(u)$ pour vecteur unitaire

$$\mathbf{a}^2(u) = 1 \text{ pour tout } u.$$

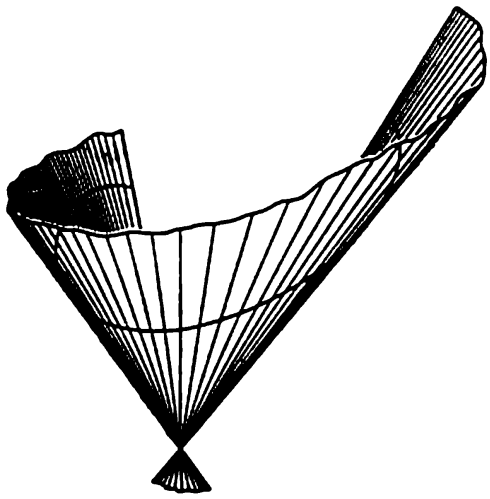


Surface réglée

Si $\rho'(u) = 0$ quel que soit u, i.e. si $\rho(u) = \text{const}$, alors l'équation (22) fait place par transport de l'origine à

$$\mathbf{r} = v\mathbf{a}(u). \tag{23}$$

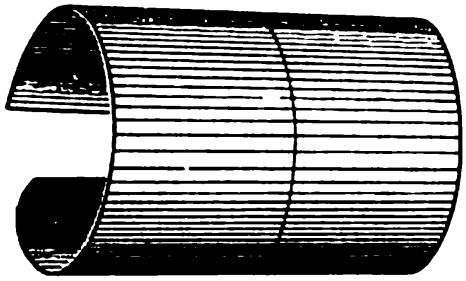
C'est un cône dont la directrice est une courbe gauche régulière $\mathbf{a} = \mathbf{a}(u)$.



Cône

Si $\mathbf{a}'(u) = 0$ quel que soit u, i.e. si $\mathbf{a}(u) = \text{const}$, alors la surface (22) est un cylindre de directrice (en général gauche) $\rho = \rho(u)$.

Dans le cas $\rho' \not\equiv 0$, on passe, si besoin est, à un domaine plus petit de $\mathbb{R}^2$, et on estime que $\rho'(u) \neq 0$ pour tout u. Dans ce cas, $\rho = \rho(u)$ est une courbe régulière dans l'espace, et on estime que u est le paramètre naturel associé (la longueur de l'arc de courbe). Le cône (23) se définit également par une équation (22), avec



Cylindre

$\boldsymbol{\rho}'(u) \neq 0$. Il suffit de poser $\boldsymbol{\rho}(u) = \mathbf{a}(u)$ dans (22) (à condition bien sûr d'avoir $\mathbf{a}'(u) \neq 0$).

Si $\mathbf{a}(u)$ est un vecteur tangent $\boldsymbol{\tau}(u)$ à la courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$, alors la surface (22) s'appelle la *surface des tangentes*. On définit de même la *surface des normales principales* et la *surface des binormales*.

S'agissant de $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ plane, la surface des binormales est un cylindre sur la courbe.

LEÇON 25

Vecteurs tangents à une surface. — Plan tangent. — Première forme quadratique d'une surface. — Longueurs et angles sur une surface. — Difféomorphismes de surfaces. — Isométrie et géométrie intrinsèque d'une surface. — Exemples. — Surfaces développables.

Par analogie avec la définition 5 de la leçon 23, on appelle *vecteur tangent* à une surface (régulière)

$$\mathbf{r} = \mathbf{r}(u, v), \quad (u, v) \in W \subset \mathbb{R}^2, \tag{1}$$

au point (u_0, v_0) un vecteur tangent à une courbe quelconque située sur la surface et passant par (u_0, v_0). Localement, une surface est le graphe d'une fonction différentiable, si bien qu'on retrouve en fait la définition mentionnée (i.e. on obtient les mêmes vecteurs). Aussi *l'ensemble de tous les vecteurs tangents à la surface* (1) *en* (u_0, v_0) *donné est un espace vectoriel de dimension* 2 (leçon 23, prop. 1). □

D'ailleurs, ce résultat est établi de façon immédiate. En effet, toute courbe de la surface (1) qui passe pour $t = t_0$ par (u_0, v_0) est définie comme courbe dans l'espace par la fonction vectorielle

$$\mathbf{r}(t) = \mathbf{r}(u(t), v(t)), \quad t_1 \leqslant t \leqslant t_2, \tag{2}$$

où $u = u(t)$ et $v = v(t)$ sont des fonctions différentiables telles que $u(t_0) = u_0$ et $v(t_0) = v_0$. Aussi

$$\mathbf{r}'(t) = u'(t)\,\mathbf{r}_u + v'(t)\,\mathbf{r}_v, \tag{3}$$

et, en particulier,

$$\mathbf{r}'(t_0) = u'(t_0)\,(\mathbf{r}_u)_0 + v'(t_0)\,(\mathbf{r}_v)_0.$$

Ainsi, tout vecteur tangent à (1) au point (u_0, v_0) est combinaison linéaire des vecteurs $(\mathbf{r}_u)_0$ et $(\mathbf{r}_v)_0$ (non colinéaires par hypothèse). Inversement, si

$$\mathbf{c} = a\,(\mathbf{r}_u)_0 + b\,(\mathbf{r}_v)_0, \tag{4}$$

avec a et b des nombres quelconques, alors $\mathbf{c} = \mathbf{r}'(t_0)$, où $\mathbf{r}(t) = \mathbf{r}(u_0 + a(t - t_0),\ v_0 + b(t - t_0))$, si bien que $\mathbf{c}$ est un vecteur tangent à (1) au point (u_0, v_0).

Du moment que les vecteurs (4) forment un espace vectoriel de dimension 2, on a le résultat voulu. □

* * *

Définition 1. L'espace vectoriel constitué de vecteurs (4) s'appelle le *plan tangent* à la surface (1) au point (u_0, v_0).

On donne le même nom au plan correspondant dans l'espace, qui passe par le point $\mathbf{r}(u_0, v_0)$. Ce plan étant de bivecteur directeur $(\mathbf{r}_u)_0 \wedge (\mathbf{r}_v)_0$ est défini en coordonnées x, y, z par l'équation

$$\begin{vmatrix} x - x(u_0, v_0) & y - y(u_0, v_0) & z - z(u_0, v_0) \\ x_u(u_0, v_0) & y_u(u_0, v_0) & z_u(u_0, v_0) \\ x_v(u_0, v_0) & y_v(u_0, v_0) & z_v(u_0, v_0) \end{vmatrix} = 0.$$

Le fait de donner un même nom à deux objets mathématiques différents n'allège certes pas l'exposé, mais il suffit d'être un peu attentif pour qu'il ne conduise à aucune ambiguïté.

Selon la formule (3), *les vecteurs* $\mathbf{r}_u$ *et* $\mathbf{r}_v$ *forment un repère du plan tangent au point* (u, v). Une tradition empruntée à l'Analyse veut qu'on note du et dv les coordonnées des vecteurs tangents par rapport à ce repère et que le vecteur de coordonnées du et dv soit désigné par $d\mathbf{r}$. Comme dans le cas de fonctions numériques, on a en multipliant à droite par les facteurs numériques:

$$d\mathbf{r} = \mathbf{r}_u\, du + \mathbf{r}_v\, dv. \tag{5}$$

Soit

$$\mathbf{r} = \mathbf{r}_1(u_1, v_1), \quad (u_1, v_1) \in W_1, \tag{6}$$

une surface équivalente à (1) (voir leçon 24, déf. 3) et soit

$$u_1 = u_1(u, v), \quad v_1 = v_1(u, v) \tag{7}$$

le difféomorphisme $W \to W_1$ associé. On a

$$\mathbf{r}(u, v) = \mathbf{r}_1(u_1(u, v), v_1(u, v))$$

pour tout point $(u, v) \in W$, si bien que

$$\begin{aligned} \mathbf{r}_u &= \frac{\partial u_1}{\partial u}(\mathbf{r}_1)_{u_1} + \frac{\partial v_1}{\partial u}(\mathbf{r}_1)_{v_1}, \\ \mathbf{r}_v &= \frac{\partial u_1}{\partial v}(\mathbf{r}_1)_{u_1} + \frac{\partial v_1}{\partial v}(\mathbf{r}_1)_{v_1}. \end{aligned} \tag{8}$$

Il en résulte que l'enveloppe linéaire des vecteurs $\mathbf{r}_u$ et $\mathbf{r}_v$ est égale à celle des $(\mathbf{r}_1)_{u_1}$ et $(\mathbf{r}_1)_{v_1}$, i.e. le plan tangent à la surface (1) au point (u, v) coïncide avec le plan tangent à (6) au point $(u_1, v_1) = (u_1(u, v), v_1(u, v))$ (qui étant un point de l'espace se confond avec (u, v)). En ce sens, *deux surfaces équivalentes ont même plan tangent.* □

Le fait de passer à une surface équivalente n'entraîne pour les plans tangents que le changement de repère (on passe de $\mathbf{r}_u$, $\mathbf{r}_v$ à $(\mathbf{r}_1)_{u_1}$, $(\mathbf{r}_1)_{v_1}$). Selon les formules (8), la matrice de passage (plus pré-

cisément, la matrice attachée au changement inverse) s'écrit

$$\begin{pmatrix} \frac{\partial u_1}{\partial u} & \frac{\partial u_1}{\partial v} \\ \frac{\partial v_1}{\partial u} & \frac{\partial v_1}{\partial v} \end{pmatrix}. \tag{9}$$

i.e. c'est la matrice jacobienne associée au difféomorphisme (7).

On en déduit, en particulier, qu'il y a entre les coordonnées du_1, dv_1 des vecteurs tangents dans le repère $(\mathbf{r}_1)_{u_1}$, $(\mathbf{r}_1)_{v_1}$ et les coordonnées du, dv de ces vecteurs dans le repère $\mathbf{r}_u$, $\mathbf{r}_v$ les relations

$$\begin{aligned} du_1 &= \frac{\partial u_1}{\partial u}\, du + \frac{\partial u_1}{\partial v}\, dv, \\ dv_1 &= \frac{\partial v_1}{\partial u}\, du + \frac{\partial v_1}{\partial v}\, dv, \end{aligned} \tag{10}$$

qui ne sont autres que les formules d'Analyse pour les différentielles des fonctions (7) (on comprend maintenant pourquoi on note du, dv les coordonnées des vecteurs tangents).

* * *

On note que le plan tangent étant un plan de l'espace euclidien est un espace euclidien de dimension 2. Depuis Gauss, les coefficients métriques g_{11}, g_{12}, g_{22} de son repère $\mathbf{r}_u$, $\mathbf{r}_v$ sont désignés par les symboles respectifs E, F et G. Ainsi, on a par définition

$$E = \mathbf{r}_u^2, \quad F = \mathbf{r}_u \mathbf{r}_v, \quad G = \mathbf{r}_v^2. \tag{11}$$

On souligne que *ces formules définissent les coefficients* $E = E(u, v)$, $F = F(u, v)$, $G = G(u, v)$ *en tant que fonctions de u et v* (d'ailleurs, c'est tout naturel car le plan tangent varie avec u, v, ainsi que son repère $\mathbf{r}_u$, $\mathbf{r}_v$). □

Définition 2. La forme quadratique

$$E\, du^2 + 2F\, du\, dv + G\, dv^2$$

des coordonnées du, dv des vecteurs tangents par rapport à $\mathbf{r}_u$, $\mathbf{r}_v$ s'appelle *la première forme quadratique* de la surface (1) et se note I. La valeur $I(d\mathbf{r})$ de I sur les coordonnées du, dv du vecteur tangent $d\mathbf{r}$ est égale au carré scalaire

$$d\mathbf{r}^2 = I(d\mathbf{r}) = E\, du^2 + 2F\, du\, dv + G\, dv^2. \tag{12}$$

Autrement dit, la forme I représente dans $\mathbf{r}_u$, $\mathbf{r}_v$ la fonctionnelle quadratique $d\mathbf{r} \longmapsto d\mathbf{r}^2$.

La première forme quadratique I_1 de la surface équivalente (6) représente donc la fonctionnelle $d\mathbf{r} \longmapsto d\mathbf{r}^2$ dans le repère $(\mathbf{r}_1)_{u_1}$, $(\mathbf{r}_1)_{v_1}$, et on obtient I par substitution dans I_1 les coordonnées du, dv données par (10).

S'agissant des coefficients E_1, F_1, G_1 de I_1, cela signifie qu'ils sont reliés à leurs homologues de la forme I par les formules

$$E(u, v) = E_1(u_1, v_1)\left(\frac{\partial u_1}{\partial u}\right)^2 + 2F_1(u_1, v_1)\frac{\partial u_1}{\partial u}\frac{\partial v_1}{\partial u} + \\ + G_1(u_1, v_1)\left(\frac{\partial v_1}{\partial u}\right)^2,$$

$$(13) \qquad F(u, v) = E_1(u_1, v_1)\frac{\partial u_1}{\partial u}\frac{\partial u_1}{\partial v} + F_1(u_1, v_1) \times \\ \times \left(\frac{\partial u_1}{\partial u}\frac{\partial v_1}{\partial v} + \frac{\partial u_1}{\partial v}\frac{\partial v_1}{\partial u}\right) + G_1(u_1, v_1)\frac{\partial v_1}{\partial u}\frac{\partial v_1}{\partial v},$$

$$G(u, v) = E_1(u_1, v_1)\left(\frac{\partial u_1}{\partial v}\right)^2 + 2F_1(u_1, v_1)\frac{\partial u_1}{\partial v}\frac{\partial v_1}{\partial v} + \\ + G_1(u_1, v_1)\left(\frac{\partial v_1}{\partial v}\right)^2.$$

Remarque. Les formules (13) s'obtiennent directement. Il suffit de porter dans (11) les valeurs (8) des vecteurs $\mathbf{r}_u$ et $\mathbf{r}_v$.

On dirait moins rigoureusement que dans le changement de coordonnées sur la surface, sa première forme quadratique subit une transformation linéaire de matrice jacobienne (9).

En d'autres termes, *les premières formes quadratiques de deux surfaces équivalentes sont équivalentes* (*en chaque point*). □

* * *

La formule (12) entraîne pour le vecteur tangent (3) à une courbe (2) quelconque de la surface (1) que sa longueur $|\mathbf{r}'(t)|$ satisfait aux égalités

$$|\mathbf{r}'(t)| = \sqrt{I(\mathbf{r}'(t))} = \sqrt{E(t)\,u'(t)^2 + 2F(t)\,u'(t)\,v'(t) + G(t)\,v'(t)^2},$$

où

$$E(t) = E(u(t), v(t)),$$
$$F(t) = F(u(t), v(t)),$$
$$G(t) = G(u(t), v(t)).$$

Comme la longueur s de la courbe (2) comprise entre les points $t = a$ et $t = b$ s'exprime par

$$s = \int_a^b |\mathbf{r}'(t)|\,dt$$

(leçon 23, (12)), on a

$$(14) \qquad s = \int_a^b \sqrt{E(t)\,u'(t)^2 + 2F(t)\,u'(t)\,v'(t) + G(t)\,v'(t)^2}\,dt,$$

ou encore en notations conventionnelles, mais plus faciles à retenir

$$s = \int_L \sqrt{E\,du^2 + 2F\,du\,dv + G\,dv^2},$$

$$s = \int_L \sqrt{I(d\mathbf{r})}.$$

La lettre L symbolise la courbe (2).

On appelle *angle* de deux courbes dans l'espace $\mathbf{r} = \mathbf{r}(t)$ et $\mathbf{r} = \mathbf{r}_1(t)$ sécantes pour t donnée l'angle φ de leurs vecteurs tangents $\mathbf{r}' = \mathbf{r}'(t)$ et $\mathbf{r}_1' = \mathbf{r}_1'(t)$. Par conséquent,

$$\cos\varphi = \frac{\mathbf{r}'\mathbf{r}_1'}{\sqrt{\mathbf{r}'^2}\sqrt{\mathbf{r}_1'^2}}$$

Si les courbes sont sur la surface (1), i.e. si

$$\mathbf{r}(t) = \mathbf{r}(u(t),\ v(t)), \qquad \mathbf{r}_1(t) = \mathbf{r}(u_1(t),\ v_1(t)),$$

la dernière formule se récrit

$$\cos\varphi = \frac{Eu'u_1' + F(u'v_1' + v'u_1') + Gv'v_1'}{\sqrt{Eu'^2 + 2Fu'v' + Gv'^2}\sqrt{Eu_1'^2 + 2Fu_1'v_1' + Gv_1'^2}}. \tag{15}$$

On pose

$$du = u'(t)\,dt, \qquad dv = v'(t)\,dt,$$
$$\delta u = u_1'(t)\,dt, \qquad \delta v = v_1'(t)\,dt,$$

il vient la formule conventionnelle

$$\cos\varphi = \frac{E\,du\,\delta u + F(du'\delta v + dv\,\delta u) + G\,dv\,\delta v}{\sqrt{E\,du^2 + 2F\,du\,dv + G\,dv^2}\sqrt{E\,\delta u^2 + 2F\,\delta u\,\delta v + G\,\delta v^2}}$$

ou, en bref,

$$\cos\varphi = \frac{d\mathbf{r}\,\delta\mathbf{r}}{\sqrt{d\mathbf{r}^2}\sqrt{\delta\mathbf{r}^2}}.$$

On la met des fois sous la forme commode

$$\cos\varphi = \frac{I(d,\ \delta)}{\sqrt{I(d)}\sqrt{I(\delta)}}.$$

En particulier, on définit le cosinus de l'angle des lignes de coordonnées $u = \text{const}$ et $v = \text{const}$ par la formule

$$\cos\varphi = \frac{F}{\sqrt{E}\sqrt{G}}.$$

Par conséquent, *les lignes de coordonnées* $u = \text{const}$ *et* $v = \text{const}$ *sont orthogonales si et seulement si* $F = 0$. □

* * *

Soit (6) une surface (régulière) arbitraire qui n'est en général pas équivalente à (1) et soit donnée une application du support de (1) dans le support de (6). Si les surfaces (1) et (6) constituent des injections (on sait qu'elles le sont localement toujours, i.e. pour W et W_1 suffisamment petits), cette application définit une application $W \to W_1$, et, inversement, toute application $W \to W_1$ définit une application du support de la surface (1) dans celui de la surface (6). A chaque application $W \to W_1$ on donne donc le nom d'*application de la surface* (1) *dans la surface* (6).

Chaque application de (1) dans (6) est donnée par définition au moyen de deux fonctions

$$u_1 = u_1(u, v), \quad v_1 = v_1(u, v) \tag{16}$$

définies pour $(u, v) \in W$ telles que $(u_1(u, v), v_1(u, v)) \in W_1$ quel que soit le point $(u, v) \in W$.

L'application (16) constitue un *difféomorphisme* de la surface (1) sur la surface (6) si elle est un difféomorphisme de l'ouvert W sur l'ouvert W_1.

Chose à souligner, *deux fonctions* (16) *non identiques définissent l'application identique de supports.* Cela a évidemment lieu si et seulement si

$$\mathbf{r}(u, v) = \mathbf{r}_1(u_1(u, v), v_1(u, v))$$

pour tout point $(u, v) \in W$, i.e. *si les fonctions* (16) *déterminent l'équivalence des surfaces* (1) *et* (6) (voir leçon 24, déf. 3). □

D'autre part, chaque fois qu'on se donne un difféomorphisme (16), on passe de la surface (6) à la surface équivalente

$$\mathbf{r} = \mathbf{r}_1(u_1(u, v), v_1(u, v)), \tag{17}$$

auquel cas la même application de supports est donnée par le difféomorphisme identique $W \to W$. Si l'on emploie un vocabulaire plus coutumier, mais moins rigoureux, on dit que (avec un choix convenable des coordonnées sur les surfaces) *tout difféomorphisme de surfaces est une application définie par égalité des coordonnées.* □

* * *

Définition 3. Le difféomorphisme (16) de la surface (1) sur la surface (6) s'appelle *isométrie* si la première forme quadratique de la surface (17) coïncide en tout point (u, v) avec la première forme quadratique de (1), i.e. si les coefficients des premières formes quadratiques des surfaces (1) et (6) vérifient les égalités (13).

Deux surfaces sont dites *isométriques* s'il existe au moins une isométrie de l'une d'elles sur l'autre.

On éclaircit cette définition par recours à une courbe L quelconque de la surface (1). Soient une fois de plus

$$u = u(t), \quad v \neq v(t), \quad a \leqslant t \leqslant b,$$

les équations paramétriques de L (considérée comme courbe sur une surface). Chaque application (16) associe à L une courbe L_1 sur la surface (6), dont les équations paramétriques s'écrivent

$$u_1 = u_1 (u (t), v (t)), \quad v_1 = v_1 (u (t), v (t)), \quad a \leqslant t \leqslant b.$$

Le support de L_1 est évidemment l'image du support de L dans l'application de supports définie par (16). D'où le nom d'*image* de L par (16) qu'on donne à la courbe L_1.

La courbe L_1 est définie sur la surface équivalente (17) par les mêmes fonctions $u = u (t)$, $v = v (t)$ que la courbe L sur (1). Aussi, la formule de la longueur s_1 de L_1 s'écrit

$$s_1 = \int_a^b \sqrt{E^* u'^2 + 2F^* u'v' + G^* v'^2}\, dt,$$

E^*, F^*, G^* étant les coefficients de la première forme quadratique de la surface (17). Si $E^* = E$, $F^* = F$ et $G^* = G$, i.e. si le difféomorphisme (16) est une isométrie, on retrouve la formule (14) pour la longueur s de la courbe L. Aussi $s_1 = s$.

Inversement, supposons qu'étant donnée une courbe L quelconque sur la surface (1), la longueur s_1 de son image L_1 sur (6) (i.e. sur la surface équivalente (17), ce qui revient au même) est égale à la fongueur s de L. Cela est vrai en particulier pour L définie par les lonctions

$$u (t) = u_0 + at, \quad v (t) = v_0 + bt, \quad 0 \leqslant t \leqslant T,$$

(u_0, v_0) étant un point arbitraire de W, a et b des nombres quelconques (et T un nombre tel que $(u (t), v (t)) \in W$ pour $0 \leqslant t \leqslant T$). Mais on a $u'(t) = a$ et $v'(t) = b$ pour ces fonctions, si bien que $s = s_1$ s'écrit

$$\int_0^T \sqrt{Ea^2 + 2Fab + Gb^2}\, dt = \int_0^T \sqrt{E^* a^2 + 2F^* ab + G^* b^2}\, dt,$$

d'où l'on tire en dérivant par rapport à T en en faisant la substitution $T = 0$

$$\sqrt{E (u_0, v_0) a^2 + 2F (u_0, v_0) ab + G (u_0, v_0) b_2} =$$
$$= \sqrt{E^* (u_0, v_0) a^2 + 2F^* (u_0, v_0) ab + G^* (u_0, v_0) b^2}.$$

Les nombres a et b étant absolument quelconques, cela ne peut avoir lieu que si

$$E (u_0, v_0) = E^* (u_0, v_0), \quad F (u_0, v_0) = F^* (u_0, v_0),$$
$$G (u_0, v_0) = G^* (u_0, v_0),$$

i.e. si $E = E^*$, $F = F^*$ et $G = G^*$ sur W (en effet, (u_0, v_0) est un point arbitraire de W), donc si le difféomorphisme (16) est une isométrie.

Ainsi, *les difféomorphismes de surfaces sont des isométries si et seulement s'ils conservent les longueurs des courbes, i.e. si toute courbe L sur la surface* (1) *et son image L_1 sur la surface* (6) *sont de même longueur.* □

Imaginons-nous une surface faite en matière élastique inextensible. La longueur des courbes situées sur la surface est alors conservée par toute déformation de celle-ci, i.e. la déformation donne une surface isométrique. Partant de cette intuition, les mathématiciens du XIXe siècle qui ont posé les fondements de la théorie des surfaces disaient « surfaces applicables l'une sur l'autre (par *flexion*)» et non « surfaces isométriques». La terminologie s'est conservée en partie, mais le terme « flexion » a un sens plus restreint : c'est une isométrie liée à la transformation identique par une famille continue d'isométries. On a longtemps cru que localement, i.e. dans un voisinage suffisamment petit d'un point quelconque, toute isométrie est cette flexion-là. Mais voilà que Nikolaï Efimov, un professeur à l'Université de Moscou, le réfute dernièrement avec un contre-exemple à l'appui.

La propriété des isométries de conserver les longueurs tient à ce que dans la formule (14) de la longueur d'une courbe figurent seulement (à part les fonctions définissant la courbe) les coefficients de la première forme quadratique I. Or, c'est également la particularité de la formule (15) pour l'angle de deux courbes. On dit donc que *les isométries conservent les angles.* □

Il y a intérêt à appeler la *géométrie intrinsèque* d'une surface l'ensemble des notions et des résultats indépendants des isométries. Ainsi, les notions de longueur et d'angle relèvent de la géométrie intrinsèque.

Il est clair qu'il en est ainsi de toute notion qui est définie par la première forme quadratique seule (comme c'est le cas de la longueur et de l'angle).

Par définition, *deux surfaces obéissent à une même géométrie intrinsèque (c'est-à-dire sont isométriques) si l'on identifie leurs premières formes quadratiques par un changement de coordonnées.* Le critère est manifestement inopérant. Notre but immédiat est de le rendre efficace. On s'en occupera dès la leçon suivante, et on propose en attendant plusieurs exemples de recherche de la première forme quadratique.

* * *

Exemple 1. Le plan Oxy rapporté aux coordonnées $u = x$ et $v = y$ a pour équation paramétrique $\mathbf{r} = u\mathbf{i} + v\mathbf{j}$. Aussi $\mathbf{r}_u = \mathbf{i}$, $\mathbf{r}_v = \mathbf{j}$, donc $E = 1$, $F = 0$, $G = 1$, i.e. la première forme quadra-

tique du plan s'écrit

$$I = du^2 + dv^2. \tag{18}$$

(D'ailleurs on l'établit facilement sans calcul.)

Exemple 2. S'agissant du cylindre circulaire

$$\mathbf{r} = R \cos u \cdot \mathbf{i} + R \sin u \cdot \mathbf{j} + v \cdot \mathbf{k},$$

on a $\mathbf{r}_u = -R \sin u \cdot \mathbf{i} + R \cos u \cdot \mathbf{j}$ et $\mathbf{r}_v = \mathbf{k}$. Aussi

$$E = \mathbf{r}_u^2 = R^2, \quad F = \mathbf{r}_u \mathbf{r}_v = 0, \quad G = \mathbf{r}_v^2 = 1,$$

i.e. pour le cylindre

$$I = R^2\, du^2 + dv^2.$$

On introduit une nouvelle coordonnée $u_1 = Ru$ (et on désigne u_1 par u), ce qui transforme I en (18). Par conséquent, *le cylindre est isométrique à un plan.*

Intuitivement, c'est évident : pour appliquer le cylindre sur un plan (plus précisément, sur une portion de plan), il suffit de le couper suivant sa génératrice.

Exemple 3. Etant donnée la surface de révolution

$$\mathbf{r} = x(v) \cos u \cdot \mathbf{i} + x(v) \sin u \cdot \mathbf{j} + z(v)\, \mathbf{k},$$

on a

$$\mathbf{r}_u = -x(v) \sin u \cdot \mathbf{i} + x(v) \cos u \cdot \mathbf{j},$$
$$\mathbf{r}_v = x'(v) \cos u \cdot \mathbf{i} + x'(v) \sin u \cdot \mathbf{j} + z'(v)\, \mathbf{k}.$$

Donc,

$$E = x(v^2) \sin^2 u + x(v)^2 \cos^2 u = x(v)^2,$$
$$F = -x(v) \sin u \cdot x'(v) \cos u + x(v) \cos u \cdot x'(v) \sin u = 0,$$
$$G = x'(v)^2 \cos^2 u + x'(v)^2 \sin^2 u + z'(v)^2 = x'(v)^2 + z'(v)^2,$$

si bien que

$$I = x(v)^2\, du^2 + (x'(v)^2 + z'(v)^2)\, dv^2.$$

On conçoit intuitivement que les méridiens et les parallèles de toute surface de révolution sont orthogonaux. Aussi l'égalité $F = 0$ s'impose sans qu'on ait besoin de calculs.

Si le profil $x = x(v)$, $z = z(v)$ de la surface de révolution est rapporté au paramètre naturel $v = s$ (auquel cas $x'(v)^2 + z'(v)^2 =$

$= 1$), la forme I devient particulièrement simple :

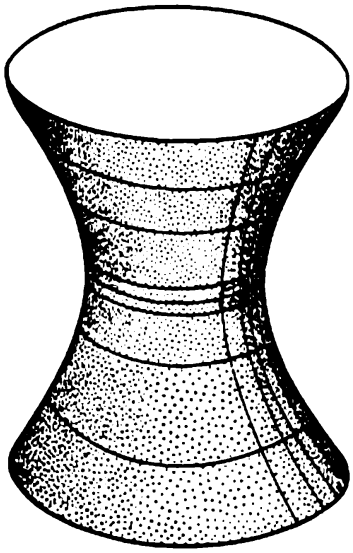
$$I = x(v)^2\, du^2 + dv^2.$$

En particulier, on a pour la sphère (de rayon 1)

$$I = \cos^2 v\, du^2 + dv^2. \tag{19}$$

Les cartographes ne savent que trop bien qu'aucune portion de sphère, fût-elle minime, n'est applicable sur le plan. Autrement dit, aucune transformation de coordonnées ne change la forme (19) en la forme (18). (On le démontrera dans la leçon terminale.)

Exemple 4. On appelle *chaînette* la courbe que prend un fil homogène qui pend sous l'action de son poids, et la surface de révolution dont le profil est la chaînette porte le nom de *caténoïde*.



Caténoïde

On montre en statique que la chaînette est le graphe du cosinus hyperbolique. Ainsi, $x(v) = \operatorname{ch} v$, $z(v) = v$ pour la caténoïde, si bien que

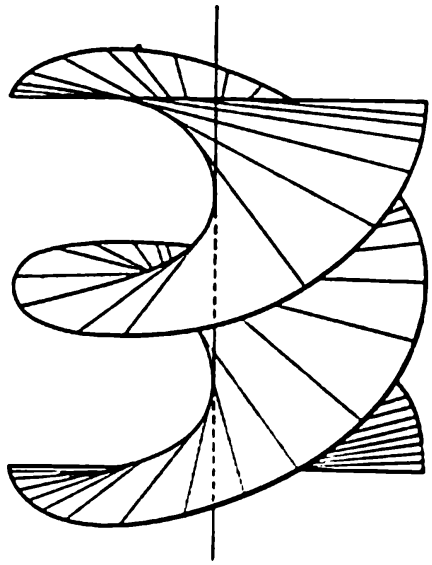
$$x(v)^2 = \operatorname{ch}^2 v \quad \text{et} \quad x'(v)^2 + z'(v)^2 = \operatorname{sh}^2 v + 1 = \operatorname{ch}^2 v.$$

La première forme quadratique de la caténoïde s'écrit donc

$$I = \operatorname{ch}^2 v\,(du^2 + dv^2). \tag{20}$$

Exemple 5. Soit une droite perpendiculaire à l'axe Oz. On suppose qu'elle subit un déplacement hélicoïdal uniforme autour de l'axe (le pas d'hélice étant proportionnel à l'angle de rotation). La

surface réglée engendrée par cette droite s'appelle l'*hélicoïde*. Elle fait penser au plan incliné des garages d'automobiles.



Hélicoïde

Si la droite est paramétrée par v, et u est l'angle de rotation, l'hélicoïde est défini par l'équation

$$\mathbf{r} = v \cos u \cdot \mathbf{i} + v \sin u \cdot \mathbf{j} + u\mathbf{k}.$$

Aussi

$$\mathbf{r}_u = -v \sin u \cdot \mathbf{i} + v \cos u \cdot \mathbf{j} + \mathbf{k},$$
$$\mathbf{r}_v = \cos u \cdot \mathbf{i} + \sin u \cdot \mathbf{j},$$

donc

$$E = 1 + v^2, \quad F = 0, \quad G = 1.$$

Ainsi, la première forme quadratique s'écrit

$$I = (1 + v^2)\, du^2 + dv^2.$$

On introduit les coordonnées u_1, v_1 reliées aux coordonnées anciennes u, v par les égalités

$$u = u_1, \quad v = \operatorname{sh} v_1.$$

Dans ce cas

$$1 + v^2 = 1 + \operatorname{sh}^2 v_1 = \operatorname{ch}^2 v_1,$$
$$du = du_1, \quad dv = \operatorname{ch} v_1\, dv_1,$$

et

$$I = \operatorname{ch}^2 v\, (du^2 + dv^2)$$

(on omet les indices des nouvelles coordonnées), forme qui coïncide avec la forme (20).

Ainsi, *l'hélicoïde et la caténoïde sont isométriques* (localement, cela va de soi), et il existe une isométrie qui transforme les méridiens de la caténoïde en les génératrices rectilignes de l'hélicoïde.

N'est-ce pas un résultat remarquable?!

Exemple 6. Soit une surface réglée arbitraire

$$\mathbf{r} = \boldsymbol{\rho}(u) + v\mathbf{a}(u), \tag{21}$$

où (voir leçon 24) $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ est une courbe régulière rapportée au paramètre naturel et $\mathbf{a}(u)$ une fonction vectorielle telle que $|\mathbf{a}(u)| = 1$ pour tout u. On dérive par rapport à u (on le symbolise par les points au-dessus des lettres), il vient

$$\mathbf{r}_u = \dot{\boldsymbol{\rho}} + v\dot{\mathbf{a}}, \quad \mathbf{r}_v = \mathbf{a}.$$

Comme $\dot{\boldsymbol{\rho}}^2 = 1$, $\dot{\mathbf{a}}\mathbf{a} = 0$ et $\mathbf{a}^2 = 1$, on a

$$E = 1 + 2v\ddot{\boldsymbol{\rho}}\mathbf{a} + v^2\dot{\mathbf{a}}^2, \quad F = \dot{\boldsymbol{\rho}}\mathbf{a}, \quad G = 1.$$

Si on a, en particulier, $\mathbf{a} = \dot{\boldsymbol{\rho}}$ (surface des tangentes), alors $\dot{\boldsymbol{\rho}}\mathbf{a} = \mathbf{a}^2 = 1$ (i.e. $F = 1$) et $\ddot{\boldsymbol{\rho}}\mathbf{a} = 0$, $\dot{\mathbf{a}}^2 = k^2$, avec k la courbure de la courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ (i.e. $E = (1 + k^2v^2)$). Ainsi, on a pour la surface des tangentes

$$I = (1 + k^2v^2)\,du^2 + 2du\,dv + dv^2. \tag{22}$$

Si $\mathbf{a}(u)$ est le vecteur de la binormale à la courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$, alors $\dot{\boldsymbol{\rho}}\mathbf{a} = 0$, $\ddot{\boldsymbol{\rho}}\mathbf{a} = 0$ et $\dot{\mathbf{a}}^2 = \varkappa^2$, où $\varkappa$ est la torsion de la courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$. Par conséquent, la première forme quadratique de la surface des binormales s'écrit

$$I = (1 + \varkappa^2v^2)\,du^2 + dv^2.$$

On voit que la première forme quadratique de la surface des tangentes dépend seulement de la courbure de la courbe, tandis que celle de la surface des binormales est fonction de la torsion seule.

* * *

On en déduit en ce qui concerne la surface des tangentes que *toute surface des tangentes est isométrique* (localement) *à un plan.* En effet, soit une courbe plane à même courbure $k = k(u)$ (cette courbe existe en vertu du théorème général 1 de la leçon 24). La première forme quadratique de la surface des tangentes à la courbe est encore (22). Il est néanmoins clair que la surface des tangentes à une courbe plane est (localement) un plan. On trouve donc un changement de coordonnées tel que la première forme quadratique $dx^2 + dy^2$ du plan devienne dans cette transformation la forme (22). (Il s'écrit

$$x = x(u) + x'(u)\,v, \quad y = y(u) + y'(u)\,v,$$

$x(u)$ et $y(u)$ étant des fonctions telles que $x'(u)^2 + y'(u)^2 = 1$ et $x''(u)^2 + y''(u)^2 = k(u)^2$.) □

Cette isométrie est réalisée en appliquant la courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ sur la courbe plane par une déformation continue.

C'est la raison pour laquelle les surfaces des tangentes portent le nom de *surfaces développables* (on sous-entend qu'elles le sont sur un plan).

Dans le cas $\mathbf{a}(u) = \boldsymbol{\rho}(u)$, la surface (21) est un cône de sommet en l'origine (et la courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ est la courbe intersection du cône et de la sphère unité $|\boldsymbol{\rho}| = 1$). On a

$$\ddot{\boldsymbol{\rho}}\mathbf{a} = \dot{\boldsymbol{\rho}}^2 = 1, \quad \dot{\mathbf{a}}^2 = 1, \quad \dot{\boldsymbol{\rho}}\mathbf{a} = 0,$$

si bien que la forme I se récrit

$$I = (1 + v)^2\, du^2 + dv^2.$$

On la transforme en la forme un peu plus simple

$$I = v^2\, du^2 + dv^2 \tag{23}$$

par le changement de coordonnées $(u, v) \mapsto (u, 1 + v)$.

On introduit les nouvelles coordonnées

$$x = v \cos u, \quad y = v \sin u.$$

Alors

$$dx = -v \sin u\, du + \cos u\, dv,$$
$$dy = v \cos u\, du + \sin u\, dv,$$

donc

$$dx^2 + dy^2 = v^2\, du^2 + dv^2.$$

Ainsi, *tout cône est isométrique à un plan*, ce qui autorise à ranger les cônes parmi les surfaces développables.

On note que la forme (23) n'est rien d'autre que la première forme quadratique du plan rapporté aux coordonnées polaires $r = v$ et $\varphi = u$.

Plaçons-nous enfin dans le cas où le vecteur $\mathbf{a}(u)$ est une constante (donc $\dot{\mathbf{a}} = 0$). La surface (21) est un cylindre. On estime, sans que cela nuise à la généralité, que la directrice $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ du cylindre est une courbe dans le plan orthogonal au vecteur $\mathbf{a}$ (si bien que $\boldsymbol{\rho}\mathbf{a} = 0$ et $\dot{\boldsymbol{\rho}}\mathbf{a} = 0$). On a donc comme dans l'exemple 2 (cylindre circulaire)

$$I = du^2 + dv^2.$$

Aussi les cylindres sont également des surfaces développables.

On montrera dans la leçon suivante que *les surfaces développables* (i.e. les cylindres, les cônes et les surfaces des tangentes) *sont les seules surfaces réglées qui soient isométriques à un plan.* Il y a plus. Les surfaces développables épuisent *l'ensemble des* surfaces isométriques à un plan. Nous nous abstiendrons de la démonstration.

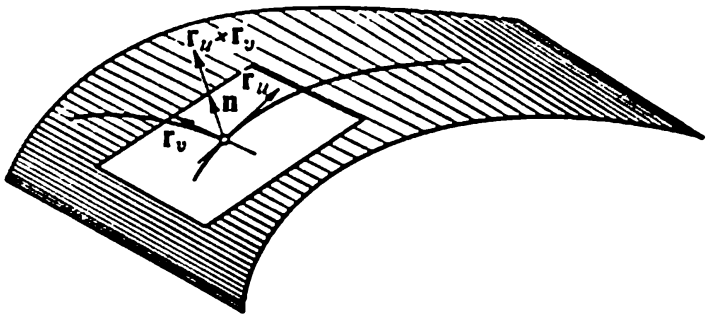
LEÇON 26

Plan tangent et vecteur de la normale. — Courbure d'une section normale. — Deuxième forme quadratique d'une surface. — Indicatrice de Dupin. — Courbures principales. — Deuxième forme quadratique d'un graphe. — Surfaces réglées à courbure nulle. — Surfaces de révolution.

Soit une fois de plus une surface régulière quelconque

(1) $$\mathbf{r} = \mathbf{r}(u, v), \quad (u, v) \in W,$$

dans l'espace euclidien $\mathscr{E}$ de dimension 3.



On rappelle que le plan tangent à (1) au point (u, v) est par définition un plan de l'espace qui passe par un point de rayon vecteur $\mathbf{r}(u, v)$ et a pour bivecteur directeur $\mathbf{r}_u \wedge \mathbf{r}_v$ (voir leçon 25). Si $\mathscr{E}$ est orienté, on définit pour tout (u, v) de la surface un vecteur unitaire $\mathbf{n} = \mathbf{n}(u, v)$ perpendiculaire au plan tangent tel que

(2) $$\mathbf{r}_u, \mathbf{r}_v, \mathbf{n}$$

constituent un repère direct de $\mathscr{E}$ (plus précisément de l'espace vectoriel attaché $\mathscr{V}$). Le vecteur $\mathbf{n}$ ainsi défini s'appelle le *vecteur de la normale* à la surface (1) au point (u, v). Le repère (2) porte le nom de *repère mobile* de la surface au point (u, v).

On souligne que le repère mobile n'est en général pas orthonormé.

Le vecteur $\mathbf{n}$ est certes colinéaire à $\mathbf{r}_u \times \mathbf{r}_v$, si bien que

$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}.$$

Lemme 1. *Deux vecteurs quelconques* $\mathbf{a}$ *et* $\mathbf{b}$ *de l'espace vectoriel euclidien orienté* $\mathscr{V}$ *de dimension* 3 *vérifient l'égalité*

$$|\mathbf{a} \times \mathbf{b}|^2 = \begin{vmatrix} \mathbf{a}^2 & \mathbf{ab} \\ \mathbf{ab} & \mathbf{b}^2 \end{vmatrix}$$

Démonstration. Soit $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ un repère orthonormé direct de l'espace vectoriel $\mathcal{V}$ tel que

$$\mathbf{a} = a\mathbf{i},$$
$$\mathbf{b} = b'\mathbf{i} + b\mathbf{j},$$

auquel cas $\mathbf{a} \times \mathbf{b} = ab\mathbf{k}$ et

$$\mathbf{a}^2 = a^2, \quad \mathbf{ab} = ab', \quad \mathbf{b}^2 = b'^2 + b^2.$$

Aussi $|\mathbf{a} \times \mathbf{b}|^2 = a^2b^2$ et

$$\begin{vmatrix} \mathbf{a}^2 & \mathbf{ab} \\ \mathbf{ab} & \mathbf{b}^2 \end{vmatrix} = \begin{vmatrix} a^2 & ab' \\ ab' & b'^2 + b^2 \end{vmatrix} = a^2(b'^2 + b^2) - (ab')^2 = a^2b^2. \qquad \square$$

Remarque. La théorie des volumes de tout espace euclidien est développée par analogie avec la théorie des aires et des volumes pour un espace de dimension 3. Le lemme constitue alors un cas particulier de la proposition générale suivante : *étant donnés les vecteurs* $\mathbf{a}_1, \ldots, \mathbf{a}_m$ *quelconques d'un espace vectoriel euclidien arbitraire* $\mathcal{V}$, *le carré du volume m-dimensionnel du parallélépipède construit sur ces vecteurs est égal au déterminant*

$$\begin{vmatrix} \mathbf{a}_1^2 & \mathbf{a}_1\mathbf{a}_2 & \ldots & \mathbf{a}_1\mathbf{a}_m \\ \mathbf{a}_2\mathbf{a}_1 & \mathbf{a}_2^2 & \ldots & \mathbf{a}_2\mathbf{a}_m \\ \cdot & \cdot & \cdot & \cdot \\ \mathbf{a}_m\mathbf{a}_1 & \mathbf{a}_m\mathbf{a}_2 & \ldots & \mathbf{a}_m^2 \end{vmatrix}.$$

Ce déterminant s'appelle le *déterminant de Gram* des vecteurs $\mathbf{a}_1, \ldots, \mathbf{a}_m$. Il est nul si et seulement si $\mathbf{a}_1, \ldots, \mathbf{a}_m$ sont linéairement dépendants. Si $m = \dim \mathcal{V}$ et $\mathbf{a}_1, \ldots, \mathbf{a}_m$ sont indépendants (i.e. ils forment une base), les éléments du déterminant de Gram ne sont autres que les coefficients métriques de la base.

S'agissant des vecteurs $\mathbf{r}_u$ et $\mathbf{r}_v$, le lemme 1 entraîne de suite

$$|\mathbf{r}_u \times \mathbf{r}_v|^2 = \begin{vmatrix} \mathbf{r}_u^2 & \mathbf{r}_u\mathbf{r}_v \\ \mathbf{r}_v\mathbf{r}_u & \mathbf{r}_v^2 \end{vmatrix} = \begin{vmatrix} E & F \\ F & G \end{vmatrix} = EG - F^2,$$

donc

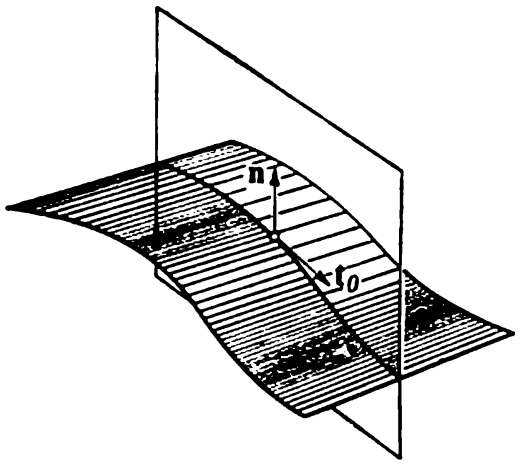
$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{\sqrt{EG - F^2}},$$

formule dont on se sert d'habitude pour calculer le vecteur $\mathbf{n}$.

* * *

Soit $\mathbf{t}_0$ un vecteur unitaire quelconque qui est un vecteur tangent à (1) au point (u_0, v_0). Considérons le plan de bivecteur directeur $\mathbf{t}_0 \wedge \mathbf{n}_0$, $\mathbf{n}_0 = \mathbf{n}(u_0, v_0)$, qui passe par un point de rayon vecteur $\mathbf{r}(u_0, v_0)$. Intuitivement, il est évident que le plan coupe la surface

suivant une courbe admettant en (u_0, v_0) le vecteur tangent $\mathbf{t}_0$ (qui est donc une courbe régulière). Cette courbe constitue la *section normale* de (1) définie par le vecteur tangent $\mathbf{t}_0$.



Section normale

Soient x, y, z des coordonnées rectangulaires dans $\mathscr{E}$ telles que la surface (1) au voisinage du point (u_0, v_0) soit le graphe d'une fonction différentiable $z = z(x, y)$, le vecteur $\mathbf{n}_0$ étant le vecteur de base $\mathbf{k}$. Si $\mathbf{t}_0 = a\mathbf{i} + b\mathbf{j}$, la section normale définie par $\mathbf{t}_0$ a évidemment (en tant que courbe sur une surface) pour équations

$$u = u_0 + at, \quad v = v_0 + bt$$

(ses équations dans l'espace sont $x = u_0 + at$, $y = v_0 + bt$, $z = z(u_0 + at, v_0 + bt)$).

On dispose donc d'un procédé pour écrire les équations d'une section normale, et, ce qui plus est, on définit formellement celle-ci (sans recourir à aucune idée intuitive) comme étant une courbe sur une surface dont les équations sont $u = u_0 + at$, $v = v_0 + bt$ (avec la condition obligatoire que la surface (1) constitue le graphe de la fonction différentiable $z = z(x, y)$). On doit certes vérifier si la définition est intrinsèque. Théoriquement, la chose n'est pas difficile, et si nous nous abstenons de le faire, c'est que la notion de section normale joue dans notre exposé un rôle subalterne, voire heuristique pour l'essentiel.

Soient $u = u(s)$, $v = v(s)$ les équations (dans le cas de surfaces) de la section normale de (1) en (u_0, v_0) définie par le vecteur tangent $\mathbf{t}_0$ et soit s le paramètre naturel d'une courbe gauche $\mathbf{r} = \mathbf{r}(s)$, où $\mathbf{r}(s) = \mathbf{r}(u(s), v(s))$, et $u(0) = u_0$, $v(0) = v_0$. Le vecteur tangent $\mathbf{t} = \mathbf{t}(s)$ à la section normale vérifie alors l'égalité

$$\mathbf{t} = \dot{\mathbf{r}} = \mathbf{r}_u \dot{u} + \mathbf{r}_v \dot{v},$$

avec $\mathbf{t}(0) = \mathbf{t}_0$. Par conséquent,

$$\dot{\mathbf{t}} = \dot{\mathbf{r}}_u \dot{u} + \mathbf{r}_u \ddot{u} + \dot{\mathbf{r}}_r \dot{v} + \mathbf{r}_v \ddot{v} =$$

$$= (\mathbf{r}_{uu}\dot{u} + \mathbf{r}_{uv}\dot{v})\,\dot{u} + (\mathbf{r}_{vu}\dot{u} + \mathbf{r}_{vv}\dot{v})\,\dot{v} + \mathbf{r}_u\ddot{u} + \mathbf{r}_v\ddot{v} =$$

$$= \mathbf{r}_{uu}(\dot{u})^2 + 2\mathbf{r}_{uv}(\dot{u}\dot{v}) + \mathbf{r}_{vv}(\dot{v})^2 + \mathbf{r}_u\ddot{u} + \mathbf{r}_v\ddot{v}.$$

On pose $s = 0$ et on multiplie par $\mathbf{n}_0$, il vient

$$(3) \qquad \dot{\mathbf{t}}(0)\,\mathbf{n}_0 = ((\mathbf{r}_{uu})_0\,\mathbf{n}_0)\,\dot{u}(0)^2 + \\ + 2((\mathbf{r}_{uv})_0\,\mathbf{n}_0)\,\dot{u}(0)\,\dot{v}(0) + ((\mathbf{r}_{vv})_0\,\mathbf{n}_0)\,\dot{v}(0)^2$$

car $(\mathbf{r}_u)_0\,\mathbf{n}_0 = 0$ et $(\mathbf{r}_v)_0\,\mathbf{n}_0 = 0$.

On note qu'une section normale est par définition une courbe plane. Les vecteurs $\mathbf{t}_0$, $\mathbf{n}_0$ définissent dans le plan de la courbe une orientation, si bien que la section normale a une courbure relative k_{rel} en chacun de ses points (voir leçon 23). Cette courbure est évidemment égale en $s = 0$ au produit scalaire $\dot{\mathbf{t}}(0)\,\mathbf{n}_0$ et est donc calculée par la formule (3)

On simplifie les formules en omettant l'indice 0, i.e. on désigne $\mathbf{t}_0$ par $\mathbf{t}$, le point (u_0, v_0) par (u, v), et ainsi de suite. La courbure relative (en $s = 0$) de la section normale en (u, v) définie par $\mathbf{t}$ sera notée $k(\mathbf{t})$. On pose de plus

$$(4) \qquad \begin{aligned} L &= \mathbf{r}_{uu}\mathbf{n} = -\mathbf{r}_u\mathbf{n}_u, \\ M &= \mathbf{r}_{uv}\mathbf{n} = -\mathbf{r}_u\mathbf{n}_v = -\mathbf{r}_v\mathbf{n}_u, \\ N &= \mathbf{r}_{vv}\mathbf{n} = -\mathbf{r}_v\mathbf{n}_v \end{aligned}$$

(comme $\mathbf{r}_u\mathbf{n} = 0$, on a $\mathbf{r}_{uu}\mathbf{n} + \mathbf{r}_u\mathbf{n}_u = 0$ et $\mathbf{r}_{uv}\mathbf{n} + \mathbf{r}_u\mathbf{n}_v = 0$, et $\mathbf{r}_v\mathbf{n} = 0$ entraîne $\mathbf{r}_{vu}\mathbf{n} + \mathbf{r}_v\mathbf{n}_u = 0$ et $\mathbf{r}_{vv}\mathbf{n} + \mathbf{r}_v\mathbf{n}_v = 0$). Avec ces notations, la formule (3) se récrit

$$(5) \qquad k(\mathbf{t}) = L\dot{u}^2 + 2M\dot{u}\dot{v} + N\dot{v}^2,$$

$\dot{u}$ et $\dot{v}$ étant les coordonnées de $\mathbf{t}$ dans le repère $\mathbf{r}_u$, $\mathbf{r}_v$:

$$\mathbf{t} = \mathbf{r}_u\dot{u} + \mathbf{r}_v\dot{v}.$$

On prend (5) pour définition formelle de la fonction $\mathbf{t} \mapsto k(\mathbf{t})$, auquel cas tous les raisonnements antérieurs constituent sa justification non formelle.

On applique la fonction construite $\mathbf{t} \mapsto k(\mathbf{t})$ à tous les vecteurs tangents non nuls $d\mathbf{r} = \mathbf{r}_u\,du + \mathbf{r}_v\,dv$, et on estime par définition que

$$k(d\mathbf{r}) = k\left(\frac{d\mathbf{r}}{ds}\right)$$

(on rappelle que $ds = |\,d\mathbf{r}\,|$; voir plus haut). Comme le vecteur unitaire $\frac{d\mathbf{r}}{ds}$ a pour coordonnées les nombres $\frac{du}{ds}$ et $\frac{dv}{ds}$, on a par la formule (5)

$$k(d\mathbf{r}) = L\left(\frac{du}{ds}\right)^2 + 2M\,\frac{du}{ds}\,\frac{dv}{ds} + N\left(\frac{dv}{ds}\right)^2 = \frac{L\,du^2 + 2M\,du\,dv + N\,dv^2}{ds^2}.$$

Puisque

$$ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2,$$

on aboutit donc à

$$k(d\mathbf{r}) = \frac{L\,du^2 + 2M\,du\,dv + N\,dv^2}{E\,du^2 + 2F\,du\,dv + G\,dv^2}. \tag{6}$$

* * *

Définition 1. La forme quadratique

$$L\,du^2 + 2M\,du\,dv + N\,dv^2$$

s'appelle la *deuxième forme quadratique* de la surface (1). On la désigne par le symbole II.

On introduit le vecteur

$$d\mathbf{n} = \mathbf{n}_u\,du + \mathbf{n}_v\,dv, \tag{7}$$

ce qui permet d'identifier (en vertu de (4)) la forme II et le produit scalaire $-d\mathbf{r}\,d\mathbf{n}$.

La formule (6) devient alors l'égalité suivante facile à retenir:

$$k = \frac{II}{I},$$

ou encore (à l'aide du vecteur (7))

$$k = -\frac{d\mathbf{r}\,d\mathbf{n}}{d\mathbf{r}^2}.$$

Certains auteurs désignent les coefficients L, M, N de la forme II par les symboles D, D', D'' respectivement.

* * *

Pour se représenter intuitivement la fonction $\mathbf{t} \mapsto k(\mathbf{t})$, on utilise une courbe du plan tangent qui porte le nom d'*indicatrice de Dupin*. On la construit pour tout vecteur unitaire $\mathbf{t}$ en menant à partir du point de tangence (pris pour origine O dans le plan tangent) dans la direction de $\mathbf{t}$ un segment de longueur $|\,k(\mathbf{t})|^{-1/2}$. On note x et y les coordonnées (dans le repère $O\mathbf{r}_u\mathbf{r}_v$) de l'extrémité du segment, auquel cas la longueur de celui-ci s'écrit avec les notations faciles à interpréter

$$|\,x\mathbf{r}_u + y\mathbf{r}_v\,| = \sqrt{I(x, y)}.$$

La formule (6) de la courbure k (t) se récrit

$$k(\mathbf{t}) = \frac{II(x, y)}{I(x, y)},$$

si bien que l'indicatrice de Dupin est définie par l'équation

$$\sqrt{I(x, y)} = \sqrt{\frac{I(x, y)}{II(x, y)}},$$

i.e. par

$$|II(x, y)| = 1.$$

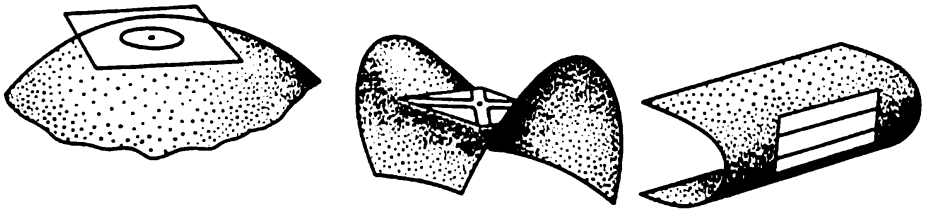
On a démontré que *l'indicatrice de Dupin est une courbe d'équation*

$$|Lx^2 + 2Mxy + Ny^2| = 1.$$

Si $LN - M^2 > 0$, cette courbe (ou, plus précisément, l'ensemble de ses points réels qui seul nous intéresse) est l'ellipse

$$Lx^2 + 2Mxy + Ny^2 = \varepsilon, \tag{8}$$

avec $\varepsilon = +1$ ou $\varepsilon = -1$ selon que $L > 0$ ou $L < 0$. Le point de la surface (1) en lequel $LN - M^2 > 0$ est dit *elliptique*.



En un point elliptique *En un point hyperbolique* *En un point parabolique*

Indicatrice de Dupin

En un point elliptique, toutes les courbures k (t) sont de même signe (qui est celui de L). Parmi celles-ci, il y en a la courbure maximale k_1 et la courbure minimale k_2 (à moins que la courbure n'est pas une constante, i.e. l'indicatrice de Dupin est une circonférence) qui correspondent aux directions du petit axe et du grand axe respectivement de l'ellipse (8).

Quand $LN - M^2 < 0$, l'indicatrice de Dupin est formée de deux hyperboles

$$Lx^2 + 2Mxy + Ny^2 = \pm 1 \tag{9}$$

admettant les mêmes asymptotes. On dit donc du point de la surface (1) en lequel $LN - M^2 < 0$ qu'il est *hyperbolique*. La courbure k (t) atteint son maximum $k_1 > 0$ dans la direction de l'axe réel de l'une des hyperboles (9). Si l'on fait tourner le vecteur t, la courbure diminue jusqu'à devenir nulle (pour t ayant la direction asymp-

totique), pour atteindre afin le minimum $k_2 < 0$ (quand la direction de t coïncide avec la direction de l'axe réel de l'autre hyperbole, i.e. avec celle de l'axe imaginaire de la première).

Si $LN - M^2 = 0$, le point correspondant de la surface (1) est dit *parabolique*. L'équation en ce point de l'indicatrice de Dupin s'écrit

$$(\sqrt{|L|}\,x + \sqrt{|N|}\,y)^2 = 1, \tag{10}$$

i.e. l'indicatrice se décompose en deux droites parallèles (sous la condition que $L \neq 0$ ou $N \neq 0$). La courbure k (t) est nulle dans la direction de ces droites, et elle est maximale (en valeur absolue) sans changer de signe dans la direction perpendiculaire. Si $L = 0$, $N = 0$ (donc $M = 0$), k (t) est identiquement nulle en t (et l'indicatrice n'est pas définie).

On note qu'en les points elliptiques et paraboliques, l'indicatrice de Dupin est une courbe du second degré, tandis qu'elle est une courbe du quatrième degré en un point hyperbolique.

Dans les trois cas, la fonction k (t) (quand elle n'est pas identiquement nulle) passe nécessairement *deux fois* par son maximum k_1 et son minimum k_2.

* * *

Définition 2. Les nombres k_1 et k_2 s'appellent les *courbures principales* de la surface (1) au point concerné. Le produit

$$K = k_1 k_2$$

est la *courbure totale* (ou *courbure de Gauss*), et la somme

$$H = \frac{k_1 + k_2}{2}$$

s'appelle la *courbure moyenne*.

Les résultats ci-dessus autorisent à dire que $K > 0$, $K < 0$ ou $K = 0$ selon que le point considéré est elliptique, hyperbolique ou parabolique.

La recherche des courbures principales pourrait consister à déterminer les directions principales des courbes du second degré (8) et (9) (le problème ne se pose pas pour la courbe (10)), puis à établir leurs équations canoniques. Les coordonnées x et y n'étant pas rectangulaires, cette manière d'agir impose des calculs laborieux. Aussi nous utiliserons la formule fondamentale (6).

Conformément à (6), la courbure k_2 est la valeur minimale de la fonction

$$\frac{II\,(x,\,y)}{I\,(x,\,y)} = \frac{Lx^2 + 2Mxy + Ny^2}{Ex^2 + 2Fxy + Gy^2}$$

de deux variables x et y pour $(x, y) \neq (0, 0)$. C'est pourquoi

$$\frac{II(x, y)}{I(x, y)} \geqslant k_2$$

pour tout $(x, y) \neq (0, 0)$, l'égalité ayant lieu au moins en un point (x, y). Du moment que $I(x, y) > 0$ pour $(x, y) \neq (0, 0)$, la dernière inégalité est équivalente à

$$II(x, y) - k_2 I(x, y) \geqslant 0$$

qui signifie que la forme quadratique $II - k_2 I$ de matrice

$$\begin{pmatrix} L - k_2 E & M - k_2 F \\ M - k_2 F & N - k_2 G \end{pmatrix}$$

est non négative en tous les points $(x, y) \neq (0, 0)$ et atteint la valeur nulle au moins en un de ces points.

De même, on caractérise le nombre k_1 par la forme quadratique $II - k_1 I$ qui est partout non positive et nulle au moins en un point $(x, y) \neq (0, 0)$.

Or on établit facilement (de façon immédiate ou moyennant la théorie générale des formes quadratiques sur $\mathbb{R}$; voir leçon 12) qu'*une forme quadratique de deux variables est partout non positive (ou partout non négative) et nulle au moins en un point* $(x, y) \neq (0, 0)$ *si et seulement si son rang est inférieur à* 2, *i.e. si sa matrice est de déterminant nul.* □

Cela prouve que *les courbures principales* k_1, k_2 *sont racines de l'équation*

$$\begin{vmatrix} L - kE & M - kF \\ M - kF & N - kG \end{vmatrix} = 0,$$

i.e. de l'équation

$$(EG - F^2)\, k^2 - (EN + GL - 2FM)\, k + (LN - M^2) = 0.$$

On en tire en particulier (en vertu des formules de Viète) que

$$K = \frac{LN - M^2}{EG - F^2}, \qquad H = \frac{1}{2}\,\frac{EN + GL - 2FM}{EG - F^2}.$$

La première formule trouvera des applications importantes dans la leçon 27.

* * *

Supposons les coordonnées x, y, z dans l'espace $\mathscr{E}$ telles que la surface considérée soit le graphe d'une fonction $z = z(x, y)$, avec $z(0, 0) = 0$, le vecteur de la normale au point $(0, 0)$ étant le vecteur unitaire $\mathbf{k}$ de l'axe Oz. On vérifie de suite que la dernière hypothèse équivaut à $\left(\frac{\partial z}{\partial x}\right)_0 = 0$ et $\left(\frac{\partial z}{\partial y}\right)_0 = 0$, si bien que le développe-

ment taylorien de $z(x, y)$ commence par les termes du second degré :

$$z = rx^2 + 2sxy + ty^2 + \ldots,$$

où

$$r = \left(\frac{\partial^2 z}{\partial x^2}\right)_0, \quad s = \left(\frac{\partial^2 z}{\partial x\, \partial y}\right)_0, \quad t = \left(\frac{\partial^2 z}{\partial y^2}\right)_0.$$

On a dans ce cas $\mathbf{r} = u\mathbf{i} + v\mathbf{j} + z(u, v)\mathbf{k}$, donc $\mathbf{r}_u = \mathbf{i} + z_u\mathbf{k}$, $\mathbf{r}_v = \mathbf{j} + z_v\mathbf{k}$ et $\mathbf{r}_{uu} = z_{uu}\mathbf{k}$, $\mathbf{r}_{uv} = z_{uv}\mathbf{k}$, $\mathbf{r}_{vv} = z_{vv}\mathbf{k}$. Par conséquent, on a au point (0, 0) les égalités $L = r$, $M = s$, $N = t$, i.e. dans ce cas, *la deuxième forme quadratique coïncide avec la somme $z_2(x, y)$ des termes du second degré du développement taylorien de la fonction $z(x, y)$.* □

Du moment que la surface $z = z(x, y)$ ne diffère guère au voisinage de (0, 0) de la surface $z = z_2(x, y)$ et que $z = z_2(x, y)$ est un paraboloïde elliptique ou un paraboloïde hyperbolique selon que $rt - s^2 > 0$ ou $rt - s^2 < 0$, on vient de prouver qu'*au voisinage d'un point elliptique une surface quelconque diffère peu d'un paraboloïde elliptique et qu'au voisinage d'un point hyperbolique elle avoisine un paraboloïde hyperbolique.* □

Cela donne une idée suffisamment nette du comportement des surfaces au voisinage des points non paraboliques.

On ne saurait dire rien de précis sur la forme des surfaces autour des points paraboliques. Généralement parlant, cette forme peut être très compliquée.

* * *

S'agissant de la surface réglée

$$\mathbf{r} = \boldsymbol{\rho}(u) + v\mathbf{a}(u), \tag{11}$$

il est déjà apparu que

$$E = 1 + 2v\dot{\boldsymbol{\rho}}\dot{\mathbf{a}} + v^2\dot{\mathbf{a}}^2, \quad F = \dot{\boldsymbol{\rho}}\mathbf{a}, \quad G = 1$$

(on fait comme toujours l'hypothèse que le paramètre u de la courbe $\boldsymbol{\rho} = \boldsymbol{\rho}(u)$ est naturel et que $\mathbf{a}(u)$ est un vecteur unitaire). Ensuite,

$$\mathbf{r}_u = \dot{\boldsymbol{\rho}} + v\dot{\mathbf{a}}, \quad \mathbf{r}_v = \mathbf{a},$$

$$\mathbf{r}_u \times \mathbf{r}_v = \dot{\boldsymbol{\rho}} \times \mathbf{a} + v(\dot{\mathbf{a}} \times \mathbf{a}),$$

$$\mathbf{n} = \frac{\dot{\boldsymbol{\rho}} \times \mathbf{a} + v(\dot{\mathbf{a}} \times \mathbf{a})}{\sqrt{EG - F^2}},$$

$$\mathbf{r}_{uu} = \ddot{\boldsymbol{\rho}} + v\ddot{\mathbf{a}}, \quad \mathbf{r}_{uv} = \dot{\mathbf{a}}, \quad \mathbf{r}_{vv} = 0,$$

$$L = \frac{(\ddot{\boldsymbol{\rho}} + v\ddot{\mathbf{a}})(\dot{\boldsymbol{\rho}} \times \mathbf{a} + v(\dot{\mathbf{a}} \times \mathbf{a}))}{\sqrt{EG - F^2}}, \quad M = \frac{\dot{\boldsymbol{\rho}}\mathbf{a}\dot{\mathbf{a}}}{\sqrt{EG - F^2}}, \quad N = 0,$$

$$LN - M^2 = -\frac{(\dot{\boldsymbol{\rho}}\mathbf{a}\dot{\mathbf{a}})^2}{EG - F^2},$$

donc

$$K = -\frac{(\dot{\rho}\mathbf{a}\dot{\mathbf{a}})^2}{(EG-F^2)^2} \leqslant 0.$$

Ainsi, *la courbure totale d'une surface réglée arbitraire est non positive, i.e. une surface réglée est dépourvue de points elliptiques.* □

Si la surface est un cylindre ($\dot{\mathbf{a}} = 0$), un cône ($\mathbf{a} = \rho$, donc $\dot{\mathbf{a}} = \dot{\rho}$) ou une surface des tangentes ($\mathbf{a} = \dot{\rho}$), la formule obtenue entraîne $K = 0$. Ainsi, *la courbure totale de chaque surface développable est nulle.*

Inversement, soit $K = 0$, auquel cas $\dot{\rho}\mathbf{a}\dot{\mathbf{a}} = 0$, i.e. les vecteurs $\dot{\rho}$, $\mathbf{a}$, $\dot{\mathbf{a}}$ sont coplanaires. Si le vecteur $\dot{\mathbf{a}}(u)$ n'est pas identiquement nul, i.e. si la surface (11) n'est pas un cylindre, alors on restreint (si besoin est) le voisinage, et on estime que $\dot{\mathbf{a}}(u) \neq 0$ pour tous les u. Aussi il y a indépendance linéaire des vecteurs $\mathbf{a}$ et $\dot{\mathbf{a}}$ ($\mathbf{a}$ et $\dot{\mathbf{a}}$ sont non nuls et orthogonaux), si bien que le vecteur $\dot{\rho}$ en est combinaison linéaire :

$$\dot{\rho} = \lambda\mathbf{a} + \mu\dot{\mathbf{a}},$$

avec $\lambda = \lambda(u)$, $\mu = \mu(u)$ des fonctions de u.

Soit

$$u_1 = u, \quad v_1 = v + \mu(u).$$

Comme le jacobien de cette transformation est 1, les nombres u_1 et v_1 sont également (après un passage éventuel à un voisinage plus petit) des coordonnées sur la surface (11), i.e. ils définissent une surface équivalente. L'équation de celle-ci est

$$\mathbf{r} = \rho_1(u_1) + v_1\mathbf{a}(u_1),$$

où

$$\rho_1(u) = \rho(u) - \mu(u)\,\mathbf{a}(u),$$

donc

$$\dot{\rho}_1 = \dot{\rho} - \dot{\mu}\mathbf{a} - \mu\dot{\mathbf{a}} = (\lambda - \dot{\mu})\mathbf{a}.$$

Si $\dot{\rho}_1$ est identiquement nul (i.e. si $\lambda = \dot{\mu}$), alors la surface (11) admet comme équation

$$\mathbf{r} = \text{const} + v_1\mathbf{a}(u_1)$$

et représente donc un cône. Dans le cas contraire, on estime (en réduisant, si besoin est, le voisinage) que $\dot{\rho}_1(u) \neq 0$ pour tout u. On passe au paramètre naturel (et on change, si besoin est, le signe de v_1), il vient $\dot{\rho}_1 = \mathbf{a}$, i.e. la surface considérée est une surface des tangentes.

On vient de démontrer la

Proposition 1. *Une surface réglée est à courbure totale nulle:*

$$K = 0,$$

si et seulement si elle est une surface développable. □

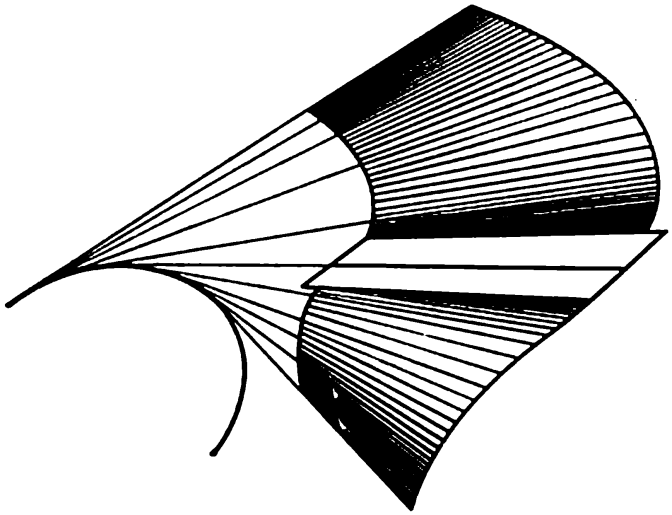
On a de plus établi que les surfaces développables sont caractérisées par la condition

$$\dot{\boldsymbol{\rho}}\mathbf{a}\dot{\mathbf{a}} = 0$$

qui équivaut, on le voit facilement, à la propriété « les vecteurs $\dot{\boldsymbol{\rho}} \times \mathbf{a}$ et $\dot{\mathbf{a}} \times \mathbf{a}$ sont colinéaires». Or, dire que $\dot{\boldsymbol{\rho}} \times \mathbf{a}$ et $\dot{\mathbf{a}} \times \mathbf{a}$ sont colinéaires, c'est dire que le vecteur

$$\mathbf{r}_u \times \mathbf{r}_v = \dot{\boldsymbol{\rho}} \times \mathbf{a} + v(\dot{\mathbf{a}} \times \mathbf{a})$$

est indépendant de v à un facteur de proportionnalité près, i.e. que le vecteur unitaire correspondant $\mathbf{n}$ est indépendant de v. Cela démontre que *les surfaces développables sont les seules surfaces réglées*



Surface des tangentes développable

à avoir un même plan tangent en tous les points d'une même génératrice rectiligne. □

* * *

Etant donnée une surface de révolution quelconque

$$\mathbf{r} = x(v)\cos u \cdot \mathbf{i} + x(v)\sin u \cdot \mathbf{j} + z(v)\,\mathbf{k},$$

on a

$$\mathbf{r}_u = -x(v)\sin u \cdot \mathbf{i} + x(v)\cos u \cdot \mathbf{j},$$
$$\mathbf{r}_v = x'(v)\cos u \cdot \mathbf{i} + x'(v)\sin u \cdot \mathbf{j} + z'(v)\,\mathbf{k}$$

donc $E = x(v)^2$, $F = 0$, $G = 1$ (on suppose que $x'(v)^2 + z'(v)^2 = 1$; voir leçon 25). Aussi

$$\mathbf{r}_u \times \mathbf{r}_v = x(v)\, z'(v) \cos u \cdot \mathbf{i} + x(v)\, z'(v) \sin u \cdot \mathbf{j} - x(v)\, x'(v)\, \mathbf{k},$$

$$\mathbf{n} = z'(v) \cos u \cdot \mathbf{i} + z'(v) \sin u \cdot \mathbf{j} - x'(v)\, \mathbf{k};$$

$$\mathbf{r}_{uu} = -x(v) \cos u \cdot \mathbf{i} - x(v) \sin u \cdot \mathbf{j},$$

$$\mathbf{r}_{uv} = -x'(v) \sin u \cdot \mathbf{i} + x'(v) \cos u \cdot \mathbf{j},$$

$$\mathbf{r}_{vv} = x''(v) \cos u \cdot \mathbf{i} + x''(v) \sin u \cdot \mathbf{j} + z''(v)\, \mathbf{k};$$

$$L = \mathbf{r}_{uu}\mathbf{n} = -x(v)\, z'(v), \qquad M = \mathbf{r}_{uv}\mathbf{n} = 0,$$

$$N = x''(v)\, z'(v) - z''(v)\, x'(v) = -\begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix},$$

$$\frac{LN - M^2}{EG - F^2} = \frac{z'(v)}{x(v)} \begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix}.$$

Cela démontre que s'agissant d'une surface de révolution

$$K = \frac{z'(v)}{x(v)} \begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix}.$$

Exemple 1. On a pour une sphère de rayon R

$$x(v) = R \cos \frac{v}{R}, \qquad z(v) = R \sin \frac{v}{R},$$

donc

$$x'(v) = -\sin \frac{v}{R}, \qquad z'(v) = \cos \frac{v}{R},$$

$$x''(v) = -\frac{1}{R} \cos \frac{v}{R} \qquad z''(v) = -\frac{1}{R} \sin \frac{v}{R},$$

$$K = \frac{z'(v)}{x(v)} \begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix} = \frac{1}{R^2}.$$

Ainsi, *une sphère de rayon R est à courbure totale constante* $\frac{1}{R^2}$. □

Intuitivement le résultat est évident.

Voici un exemple plus intéressant.

Exemple 2. Une surface de révolution de profil

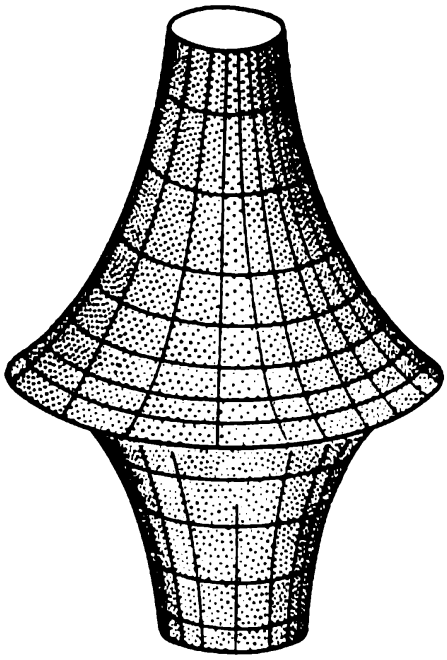
$$x(v) = R \sin v, \quad z(v) = R\left(\ln \operatorname{tg} \frac{v}{2} + \cos v\right), \quad 0 < v < \frac{\pi}{2},$$

(c'est la *tractrice*) s'appelle *pseudo-sphère*. On a pour cette surface

$$x'(v) = R \cos v, \quad z'(v) = \frac{R}{\sin v} - R \sin v = R\, \frac{\cos^2 v}{\sin v};$$

donc

$$x'(v)^2 + z'(v)^2 = R^2 \operatorname{cotg}^2 v.$$



Pseudo-sphère

Comme $x'(v)^2 + z'(v)^2 \neq 1$, on ne saurait utiliser directement la formule générale ci-dessus. On passe au préalable au paramètre naturel du profil.

On a

$$s = -R \int_{\frac{\pi}{2}}^{v} \operatorname{cotg} v \, dv = -R \ln \sin v,$$

donc

$$\sin v = e^{-\frac{s}{R}}, \quad \cos v = \sqrt{1 - e^{-2\frac{s}{R}}},$$

$$\operatorname{tg} \frac{v}{2} = e^{\frac{s}{R}} - \sqrt{e^{2\frac{s}{R}} - 1}.$$

Ainsi, la tractrice rapportée au paramètre naturel (qu'on désigne encore par v) est définie par les fonctions

$$x(v) = Re^{\frac{v}{R}},$$

$$z(v) = R \ln \left(e^{\frac{v}{R}} - \sqrt{e^{2\frac{v}{R}} - 1}\right) + R\sqrt{1 - e^{-2\frac{v}{R}}}.$$

On calcule

$$x'(v) = -e^{-\frac{v}{R}}, \quad z'(v) = -\sqrt{1-e^{-2\frac{v}{R}}},$$

$$x''(v) = \frac{1}{R}e^{-\frac{v}{R}}, \quad z''(v) = -\frac{e^{-2\frac{v}{R}}}{R\sqrt{1-e^{-2\frac{v}{R}}}},$$

$$\begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix} = -\frac{e^{-2\frac{v}{R}}}{R\sqrt{1-e^{-2\frac{v}{R}}}},$$

$$\frac{z'(v)}{x(v)}\begin{vmatrix} x''(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix} = -\frac{1}{R^2}.$$

Ainsi

$$K = -\frac{1}{R^2},$$

i.e. *la courbure totale de la pseudo-sphère est constante et vaut* $-\frac{1}{R^2}$. □

En ce qui concerne la courbure totale, la pseudo-sphère diffère de la sphère par le seul signe de la courbure (d'où le nom de pseudo-sphère).

Exemple 3. On a pour la caténoïde

$$x(v) = \operatorname{ch} v, \quad z(v) = v,$$
$$x'(v) = \operatorname{sh} v, \quad z'(v) = 1,$$
$$x'(v)^2 + z'(v)^2 = \operatorname{ch}^2 v,$$

si bien qu'on doit passer au paramètre naturel

$$s = \int_0^v \operatorname{ch} v\, dv = \operatorname{sh} v.$$

On le désigne par v, il vient

$$x(v) = \sqrt{1+v^2}, \quad z(v) = \ln(v + \sqrt{1+v^2}).$$

Aussi

$$x'(v) = \frac{v}{\sqrt{1+v^2}}, \quad z'(v) = \frac{1}{\sqrt{1+v^2}},$$

$$x''(v) = \frac{1}{(1+v^2)^{3/2}}, \quad z''(v) = -\frac{v}{(1+v^2)^{3/2}},$$

$$\begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix} = -\frac{1}{1+v^2}$$

et, partant,

$$K = -\frac{1}{(1+v^2)^2}.$$

Il y a intérêt à comparer la courbure de la caténoïde avec celle de l'hélicoïde (on rappelle que la caténoïde et l'hélicoïde sont isométriques).

On a pour l'hélicoïde l'équation (11) avec

$$\boldsymbol{\rho}(u) = u\mathbf{k}, \quad \mathbf{a}(u) = \cos u \cdot \mathbf{i} + \sin u \cdot \mathbf{j}.$$

Aussi

$$\dot{\boldsymbol{\rho}} = \mathbf{k}, \quad \dot{\mathbf{a}} = -\sin u \cdot \mathbf{i} + \cos u \cdot \mathbf{j},$$

$$E = 1 + 2v\dot{\boldsymbol{\rho}}\dot{\mathbf{a}} + v^2\dot{\mathbf{a}}^2 = 1 + v^2,$$

$$F = \dot{\boldsymbol{\rho}}\mathbf{a} = 0, \quad G = 1,$$

$$EG - F^2 = 1 + v^2,$$

$$\dot{\boldsymbol{\rho}}\mathbf{a}\dot{\mathbf{a}} = \begin{vmatrix} 0 & 0 & 1 \\ \cos u & \sin u & 1 \\ -\sin u & \cos u & 0 \end{vmatrix} = 1,$$

donc

$$K = -\frac{1}{(1+v^2)^2}.$$

On retrouve la courbure de la caténoïde ! Cela signifie que si *la caténoïde est applicable sur l'hélicoïde, ils ont même courbure totale aux points homologues.* □

Quelle est la courbure moyenne de deux dernières surfaces?

Cas de la caténoïde. On a $E = 1 + v^2$, $F = 0$, $G = 1$. En outre,

$$L = -x(v)\,z'(v) = -1, \quad M = 0,$$

$$N = -\begin{vmatrix} x'(v) & z'(v) \\ x''(v) & z''(v) \end{vmatrix} = \frac{1}{1+v^2},$$

donc

$$EN + GL - 2FM = 0,$$

i.e.

$$H = 0,$$

Ainsi, *la courbure moyenne de la caténoïde est nulle.* □

Cas de l'hélicoïde. On a

$$\dot{\boldsymbol{\rho}} \times \mathbf{a} = \sin u \cdot \mathbf{i} - \cos u \cdot \mathbf{j}, \quad \dot{\mathbf{a}} \times \mathbf{a} = -\mathbf{k},$$

$$\ddot{\boldsymbol{\rho}} = 0, \quad \ddot{\mathbf{a}} = -\cos u \cdot \mathbf{i} - \sin u \cdot \mathbf{j},$$

$$(\ddot{\boldsymbol{\rho}} + v\ddot{\mathbf{a}})(\dot{\boldsymbol{\rho}} \times \mathbf{a} + v(\dot{\mathbf{a}} \times \mathbf{a})) = 0$$

et (on l'a déjà constaté)

$$E = 1 + v^2, \quad F = 0, \quad G = 1,$$

$$EG - F^2 = 1 + v^2, \quad \dot{\rho}\dot{a}\dot{a} = 1.$$

Aussi

$$L = 0, \quad M = \frac{1}{\sqrt{1+v^2}}, \quad N = 0$$

et, partant,

$$EN + LG - 2FM = 0,$$

i.e.

$$H = 0.$$

Ainsi, *la courbure moyenne de l'hélicoïde est nulle.* □

Les derniers résultats autorisent à supposer que la déformation étudiée (l'isométrie) conserve la courbure totale et la courbure moyenne. *Cela est vrai* en ce qui concerne la courbure totale (on le montrera dans la leçon suivante), mais s'agissant de la courbure moyenne, l'*hypothèse est fausse.* En effet, un plan est à courbure moyenne nulle, tandis que la courbure moyenne d'un cylindre circulaire de rayon R applicable sur un plan est évidemment égale à $\frac{1}{2R}$.

Quant aux causes qui ont produit la propriété, pour la caténoïde et l'hélicoïde, d'avoir même courbure moyenne, elles sont très sérieuses et intéressantes. Le temps nous manque malheureusement pour les discuter.

LEÇON 27

Formules de dérivation de Weingarten. — Coefficients de connexion. — Théorème de Gauss. — Conditions nécessaires et suffisantes de surfaces isométriques.

On écrit pour le repère mobile $\mathbf{r}_u$, $\mathbf{r}_v$, $\mathbf{n}$ d'une surface quelconque

$$\mathbf{r} = \mathbf{r}(u, v) \tag{1}$$

des formules analogues aux résultats de Frenet pour les courbes. Ces formules donnent la décomposition des dérivées

$$\mathbf{r}_{uu},\ \mathbf{r}_{uv},\ \mathbf{r}_{vv},\ \mathbf{n}_u,\ \mathbf{n}_v$$

des vecteurs du repère mobile par rapport au même repère.

Vu que $\mathbf{n}^2 = 1$, donc $\mathbf{n}\mathbf{n}_u = 0$ et $\mathbf{n}\mathbf{n}_v = 0$, les vecteurs $\mathbf{n}_u$ et $\mathbf{n}_v$ se décomposent par rapport à $\mathbf{r}_u$ et $\mathbf{r}_v$ seuls, si bien que

$$\mathbf{n}_u = \alpha\mathbf{r}_u + \beta\mathbf{r}_v,$$
$$\mathbf{n}_v = \alpha_1\mathbf{r}_u + \beta_1\mathbf{r}_v.$$

On multiplie la première égalité par $\mathbf{r}_u$ et $\mathbf{r}_v$, il vient

$$-L = \mathbf{r}_u\mathbf{n}_u = \alpha\mathbf{r}_u^2 + \beta\mathbf{r}_u\mathbf{r}_v = \alpha E + \beta F,$$
$$-M = \mathbf{r}_v\mathbf{n}_u = \alpha\mathbf{r}_u\mathbf{r}_v + \beta\mathbf{r}_v^2 = \alpha F + \beta G,$$

d'où

$$\alpha = \frac{FM - GL}{EG - F^2}, \qquad \beta = \frac{FL - EM}{EG - F^2}.$$

On calcule de même les coefficients de la seconde égalité :

$$\alpha_1 = \frac{FN - GM}{EG - F^2}, \qquad \beta_1 = \frac{FM - EN}{EG - F^2}.$$

Par définition,

$$\mathbf{r}_{uu}\mathbf{n} = L, \quad \mathbf{r}_{uv}\mathbf{n} = M, \quad \mathbf{r}_{vv}\mathbf{n} = N,$$

et $\mathbf{r}_u\mathbf{n} = 0$, $\mathbf{r}_v\mathbf{n} = 0$ par hypothèse, si bien que les coefficients de $\mathbf{n}$ dans la décomposition de $\mathbf{r}_{uu}$, $\mathbf{r}_{uv}$, $\mathbf{r}_{vv}$ dans le repère $\mathbf{r}_u$, $\mathbf{r}_v$, $\mathbf{n}$ sont égaux à L, M, N respectivement.

Ainsi,

$$
(2)\qquad
\begin{aligned}
\mathbf{r}_{uu} &= \Gamma^1_{11}\mathbf{r}_u + \Gamma^2_{11}\mathbf{r}_v + L\mathbf{n},\\
\mathbf{r}_{uv} &= \Gamma^1_{12}\mathbf{r}_u + \Gamma^2_{12}\mathbf{r}_v + M\mathbf{n},\\
\mathbf{r}_{vv} &= \Gamma^1_{22}\mathbf{r}_u + \Gamma^2_{22}\mathbf{r}_v + N\mathbf{n},\\
\mathbf{n}_u &= \frac{FM-GL}{EG-F^2}\,\mathbf{r}_u + \frac{FL-EM}{EG-F^2}\,\mathbf{r}_v,\\
\mathbf{n}_v &= \frac{FN-GM}{EG-F^2}\,\mathbf{r}_u + \frac{FM-EN}{EG-F^2}\,\mathbf{r}_v,
\end{aligned}
$$

avec Γ^k_{ij}, i, j, $k = 1, 2$, des fonctions de u et v. Ces fonctions ont été naguère désignées par le symbole

$$\left\{ \begin{matrix} ij \\ k \end{matrix} \right\},$$

et on les appelait les *accolades de Christoffel*. Aujourd'hui, on leur donne d'habitude le nom de *coefficients de connexion*.

Les formules (2) s'appellent les *formules de dérivation de Weingarten*.

* * *

La recherche des coefficients de connexion Γ^k_{ij} débute par le calcul des produits des vecteurs $\mathbf{r}_{uu}$, $\mathbf{r}_{uv}$, $\mathbf{r}_{vv}$ par $\mathbf{r}_u$ et $\mathbf{r}_v$.

Comme $\mathbf{r}_u^2 = E$, on a $2\mathbf{r}_{uu}\mathbf{r}_u = E_u$ et $2\mathbf{r}_{uv}\mathbf{r}_u = E_v$, i.e.

$$\mathbf{r}_{uu}\mathbf{r}_u = \frac{1}{2}E_u \quad \text{et} \quad \mathbf{r}_{uv}\mathbf{r}_u = \frac{1}{2}E_v$$

De même, $\mathbf{r}_v^2 = G$ entraîne

$$\mathbf{r}_{uv}\mathbf{r}_v = \frac{1}{2}G_u \quad \text{et} \quad \mathbf{r}_{vv}\mathbf{r}_v = \frac{1}{2}G_v.$$

Puisque $\mathbf{r}_u\mathbf{r}_v = F$, on a $\mathbf{r}_{uu}\mathbf{r}_v + \mathbf{r}_u\mathbf{r}_{uv} = F_u$ et $\mathbf{r}_{uv}\mathbf{r}_v + \mathbf{r}_u\mathbf{r}_{vv} = F_v$, d'où

$$\mathbf{r}_{uu}\mathbf{r}_v = F_u - \frac{1}{2}E_v \quad \text{et} \quad \mathbf{r}_{vv}\mathbf{r}_u = F_v - \frac{1}{2}G_u.$$

On multiplie les premières trois formules (2) par $\mathbf{r}_u$ et $\mathbf{r}_v$, il vient

$$\begin{cases} E\Gamma^1_{11} + F\Gamma^2_{11} = \dfrac{1}{2}E_u, \\[2ex] F\Gamma^1_{11} + G\Gamma^2_{11} = F_u - \dfrac{1}{2}E_v, \end{cases}$$

$$\begin{cases} E\Gamma_{12}^1 + F\Gamma_{12}^2 = \frac{1}{2}E_v, \\ F\Gamma_{12}^1 + G\Gamma_{12}^2 = \frac{1}{2}G_u, \end{cases} \quad \begin{cases} E\Gamma_{22}^1 + F\Gamma_{22}^2 = F_v - \frac{1}{2}G_u, \\ F\Gamma_{22}^1 + G\Gamma_{22}^2 = \frac{1}{2}G_v. \end{cases}$$

Les coefficients Γ_{ij}^k s'en déduisent sans peine.

(Les équations admettent chacune une solution unique parce que le déterminant $EG - F^2$ de chaque couple de relations est non nul.)

On voit que *les coefficients de connexion Γ_{ij}^k s'expriment par les coefficients de la première forme quadratique et par leurs dérivées. Ils ne changent donc pas dans la flexion (l'isométrie) de la surface.* □

Nous n'aurons pas besoin des Γ_{ij}^k explicités par les coefficients de la première forme quadratique, si bien que nous nous abstiendrons d'écrire les formules explicites correspondantes.

* * *

Il y a entre les coefficients des formules de dérivation trois relations provenant de deux procédés de recherche des dérivées partielles $\mathbf{r}_{uuv}$, $\mathbf{r}_{uvv}$, $\mathbf{n}_{uv}$ à l'aide de ces formules. L'une des relations est due à Gauss et les deux autres ont été établies par Peterson, Meinardus et Codazzi. Nous nous bornerons au cas Gauss obtenu en calculant le coefficient de $\mathbf{r}_v$ dans la décomposition de la dérivée partielle $\mathbf{r}_{uuv}$ par rapport aux vecteurs $\mathbf{r}_u$, $\mathbf{r}_v$ et $\mathbf{n}$.

Seuls nous intéresseront en l'occurrence le coefficient de $\mathbf{r}_v$ et les termes qui dépendent des coefficients de la deuxième forme quadratique, les termes restants étant remplacés par des points de suspension.

On a

$$\begin{aligned} \mathbf{r}_{uuv} = (\mathbf{r}_{uu})_v &= (\Gamma_{11}^1\mathbf{r}_u + \Gamma_{11}^2\mathbf{r}_v + L\mathbf{n})_v = \\ &= \ldots + \Gamma_{11}^1\mathbf{r}_{uv} + \ldots + \Gamma_{11}^2\mathbf{r}_{vv} + \ldots + L\mathbf{n}_v = \\ &= \ldots + \Gamma_{11}^1(\ldots) + \ldots + \Gamma_{11}^2(\ldots) + \ldots \\ &\ldots + L\left(\ldots + \frac{FM-EN}{EG-F^2}\mathbf{r}_v\right) = \left(L\frac{FM-EN}{EG-F^2} + \ldots\right)\mathbf{r}_v + \ldots \end{aligned}$$

De même,

$$\mathbf{r}_{uuv} = (\mathbf{r}_{uv})_u = (\Gamma_{12}^1\mathbf{r}_u + \Gamma_{22}^2\mathbf{r}_v + M\mathbf{n})_u = \left(M\frac{FL-EM}{EG-F^2} + \ldots\right)\mathbf{r}_v + \ldots$$

Par conséquent,

$$L\frac{FM-EN}{EG-F^2} = M\frac{FL-EM}{EG-F^2} + \ldots,$$

les points de suspension désignant les termes dépendant des seuls coefficients de la première forme quadratique. Mais

$$M\frac{FL-EM}{EG-F^2} - L\frac{FM-EN}{EG-F^2} = E\frac{LN-M^2}{EG-F^2} = EK.$$

Du moment que $E \neq 0$ (la forme I est définie positive), cela prouve que *la courbure totale K d'une surface est représentée par les coefficients de la première forme quadratique* (et par leurs dérivées). Il en résulte que la courbure K ne change pas dans la flexion. Ce résultat mérite qu'on l'énonce sous forme de

Théorème 1 (Gauss). *Si une surface courbe est développée sur une autre surface quelconque, la courbure totale (de Gauss) en chaque point reste invariable, i.e. deux surfaces isométriques ont même courbure aux points homologues.* □

Gauss tira grand orgueil de ce théorème qu'il a qualifié de beau. Ce résultat entraîne en particulier qu'aucune partie, fût-elle minime, d'une sphère n'est applicable sur un plan, si bien que la surface terrestre n'est jamais représentée fidèlement par les cartes géographiques.

La courbure K représentée par les coefficients E, F et G de I s'écrit sous forme explicite

$$K = -\frac{1}{4(EG-F^2)^2}\begin{vmatrix} E & E_u & E_v \\ F & F_u & F_v \\ G & G_u & G_v \end{vmatrix} - \frac{1}{2\sqrt{EG-F^2}} \times \tag{3}$$
$$\times \left\{ \left(\frac{E_v - F_u}{\sqrt{EG-F^2}} \right)_v - \left(\frac{F_v - G_u}{\sqrt{EG-F^2}} \right)_u \right\}.$$

Les deux autres relations provenant des formules de dérivation sont d'ordinaire connues sous le nom de *formules de Peterson-Codazzi.* Elles s'écrivent

$$\begin{aligned} &2(EG-F^2)(L_v - M_u) - (EN + GL - 2FM)(E_v - F_u) + \\ &\qquad + \begin{vmatrix} E & E_u & L \\ F & F_u & M \\ G & G_u & N \end{vmatrix} = 0, \\ &2(EG-F^2)(M_v - N_u) - (EN + GL - 2FM)(F_v - G_u) + \\ &\qquad + \begin{vmatrix} E & E_v & L \\ F & F_v & M \\ G & G_v & N \end{vmatrix} = 0. \end{aligned} \tag{4}$$

On les démontre facilement à condition de faire preuve de la patience et d'une grande attention.

* * *

Le théorème de Gauss dit que l'égalité des courbures totales de deux surfaces est une condition nécessaire pour qu'elles soient isométriques. La condition n'est pas suffisante, mais elle est si forte

qu'on obtient avec son aide des conditions suffisantes. Nous ne nous proposons pas d'entrer dans le détail de la question car seul nous intéresse un cas particulier très important du théorème correspondant.

Soit

$$\Delta_1 K = \frac{E K_v^2 - 2F K_u K_v + G K_u^2}{EG - F^2}.$$

(C'est le premier paramètre différentiel de Beltrami de la fonction K qui est calculé en coordonnées « curvilignes» u et v.) S'il y a *indépendance fonctionnelle* des fonctions K et $\Delta_1 K$ de u et v, i.e. si leur jacobien

$$\begin{vmatrix} \dfrac{\partial K}{\partial u} & \dfrac{\partial K}{\partial v} \\ \dfrac{\partial \Delta_1 K}{\partial u} & \dfrac{\partial \Delta_1 K}{\partial v} \end{vmatrix}$$

est non nul, on prend u et v pour nouvelles coordonnées locales sur la surface (1). Nous dirons de ces coordonnées qu'elles sont *gaussiennes*. On montre par un calcul direct que dès qu'un difféomorphisme de surfaces (en particulier, une isométrie) conserve la fonction K, elle laisse invariante la fonction $\Delta_1 K$. Aussi, on dit en particulier que *toute isométrie est une application définie par égalité des coordonnées gaussiennes*. Cela signifie la validité du

Théorème 2. *Deux surfaces sur lesquelles on a défini les coordonnées gaussiennes, sont isométriques si et seulement si elles ont mêmes premières formes quadratiques en ces coordonnées.* □

Pour vérifier si oui ou non deux surfaces sont isométriques, on définit donc sur elles (si possible) les coordonnées gaussiennes et on calcule en ces coordonnées leurs premières formes quadratiques. S'il y a égalité des formes, les surfaces sont isométriques, et elles ne le sont pas dans le cas contraire.

Le théorème 2 ne résout pas le problème s'il y a dépendance fonctionnelle de K et $\Delta_1 K$, par exemple si $\Delta_1 K = 0$ (on conçoit sans peine que cette égalité a lieu si et seulement si $K = \text{const}$). S'agissant de ce cas extrême, on montre d'ailleurs que la condition du théorème 1 est suffisante, i.e. *deux surfaces à courbure totale constante sont isométriques si et seulement si elles ont même courbure*. Autrement dit, *toute surface à courbure totale K constante est isométrique à une sphère de rayon $R = \frac{1}{\sqrt{K}}$, à un plan ou à une pseudo-sphère de paramètre $R = \frac{1}{\sqrt{-K}}$ selon que $K > 0$, $K = 0$ ou $K < 0$*. On démontre ce résultat en construisant explicitement les coordonnées u et v en lesquelles la première forme quadratique coïncide avec la première forme quadratique de la sphère, du plan et de la pseudo-sphère respectivement. Une fois de plus le temps nous manque pour le faire.

INDEX

Abaissement d'un indice 152
Accolades de Christoffel 256
Algèbre de Grassmann 63
Algèbre tensorielle des espaces vectoriels 49
Algorithme de Lagrange 92
Alternative de Fredholm 118
Annulateur d'un ensemble 36
Anti-isomorphisme d'algèbres 117
Application alternée 55
Application linéaire 27

Base 10
  duale d'une base 32
Bloc de Jordan 134

Caténoïde 235
Chaînette 235
Champ vectoriel 184
  central 188
  irrotationnel 187
  de potentiel 187
  solénoïdal 190
  symbolique de Hamilton 191
Coefficients de connexion 256
Coefficients
  stricts d'une fonctionnelle antisymétrique 67
  d'un tenseur 45
Coimage d'un homomorphisme d'espaces vectoriels 28
Complexifié
  d'un espace vectoriel 137
  d'un opérateur linéaire 138
Cône multiple d'ordre k 102
Coordonnées
  contravariantes 152
  covariantes 152
  plückériennes 79
Couplage 32
Courbe dans l'espace euclidien (affine) 197
  différentiable 198
  régulière 199
  de type général 208, 209, 214
Courbes
  équivalentes 199
  non paramétrées 199
Courbure d'une courbe 207, 209, 215
  relative 207
Courbure(s) d'une surface 244
  moyenne 245
  principales 245
  totale 245
Covecteur 30
  associé à une fonctionnelle multilinéaire 51
Critère de Sylvester 100
Cylindre 221
  multiple d'ordre k 102

Degré d'une fonctionnelle multilinéaire 51
Difféomorphisme 219
Divergence d'un champ vectoriel 190
Droite 80

Elévation d'un indice 152
Enveloppe linéaire d'un ensemble 11
Epimorphisme d'espaces vectoriels 27
Equation paramétrique vectorielle du plan 80
Espace
  associé à une fonctionnelle 51

dual d'un espace vectoriel 30
pseudo-euclidien 92
vectoriel 10
Espace vectoriel quotient d'un espace vectoriel par un sous-espace 26
Espaces vectoriels en dualité 32

Fonction harmonique 195
Fonctionnelle
bilinéaire 40
— mixte 43
linéaire 30
multilinéaire 51
— antisymétrique 59
— symétrique 86
quadratique 89
— définie positive 100
sesquilinéaire 144
Forme
bilinéaire 40
normale de Jordan 134
quadratique 90
— canonique 94
— définie positive 100
Formules
de dérivation de Weingarten 256
de Frenet 209, 210, 215
de Peterson-Codazzi 258

Gradient d'une fonction 180
Graphe d'une fonction 181
Groupe
symétrique 54
unitaire 173

Hélice 210
Hélicoïdale 236
Homomorphisme d'espaces vectoriels 27
Hyperplan 80
tangent 204
Hypersurface différentiable (ou régulière) 180
Hypersurface du second degré de l'espace affine 105
à centre 106
Hypersurface du second degré de l'espace projectif 102
ovale 102

Image d'un homomorphisme d'espaces vectoriels 27
Indicatrice de Dupin 243
Indice d'inertie
négatif 99
positif 99
Isométrie de surfaces 231

Laplacien 195
Loi d'inertie des formes quadratiques 98
Loi de transformation des tenseurs 46
Longueur d'une courbe continue 205

Matrice
des coefficients 20
— allongée 21
définie positive 100
d'une fonctionnelle bilinéaire 41
d'un opérateur linéaire 114
principale extraite d'une matrice 95
unitriangulaire 95
Mineur 16
principal 95
Monomorphisme d'espaces vectoriels 27
Morphisme d'espaces vectoriels 27
Multivecteur de degré p 70

Nœud 202
Noyau d'un homomorphisme d'espaces vectoriels 27

Opérateur linéaire 112
adjoint 154
antihermitien 156
antisymétrique 156
auto-adjoint 155
cyclique 127
dégénéré 115
diagonalisable 125
hermitien 155
identique 113
inversible 115
isométrique 172
nilpotent 127
non dégénéré 115
non négatif 169
normal 166

BeatriceGloria_personal library

nul 112
orthogonal 172
orthogonalement diagonalisable 160
positif 169
scalaire 113
symétrique 155
unitaire 172
Ordre de nilpotence d'un opérateur (d'une matrice) 127

Paraboloïde 110
elliptique 110
hyperbolique 110
Paramètre différentiel
de Beltrami, premier 196
de Darboux 196
Permutation 54
impaire 54
paire 54
Plan
de dimension p 80
de dimension r de l'espace projectif 83
normal 214
osculateur 213
rectifiant 214
Point
elliptique d'une surface 244
hyperbolique d'une surface 244
parabolique d'une surface 245
singulier d'un champ vectoriel 185
— d'une hypersurface 185
Polynôme
caractéristique 122
d'un opérateur 135
Potentiel-vecteur 190
Produit extérieur 61
Produit tensoriel
de fonctionnelles linéaires 41
de vecteurs 42
Propriété de maximum-minimum des valeurs propres 165
Pseudo-sphère 250
p-vecteur directeur d'un sous-espace 79

Rang
d'un ensemble 15
extérieur 72
d'une fonctionnelle multilinéaire 53
d'une matrice 16
matriciel d'une fonctionnelle bilinéaire 88
Rayon de courbure 208
Relations de Plücker 77
Repère mobile de Frenet 208, 209, 215
Réseau de coordonnées 221
Restriction d'un opérateur linéaire 119
Rotation 175
généralisée 175
Rotationnel 188

Somme directe d'espaces 29
Somme de sous-espaces 13
directe 23
Sous-espace(s)
invariant 119
nul 11
principal 131
propre à une valeur propre 121
supplémentaires 25
trivial 11
vectoriel 10
Spectre d'un opérateur linéaire 126
simple 126
Supplémentaire orthogonal 145
Support d'une surface 219
Surface(s)
des binormales 225
deuxième forme quadratique 243
développables 238
isométriques 231
des normales principales 225
première forme quadratique 228
réglée 223
régulière 219
des tangentes 225
Symbole de Kronecker 34
Système fondamental de solutions 37-38

Tenseur contracté 53
Tenseur de type (p, q) 44
Théorème
de Gauss 258
de Hamilton-Cayley 136
de Jacobi 111
de Kronecker-Capelli 21
de Lagrange 92
Torsion 210
Trace 50
Tractrice 210
Transformation
affine 173

Transformation centrée 174
orthogonale 173
unitaire 174
Translation 174

Valeur propre 121
multiplicité algébrique 122
— géométrique 121

Variété
grassmannienne 85
linéaire 80
Vecteur 10
de la binormale 209
de la normale principale à une courbe 209
— — à une surface 239
principal 131
tangent 198, 203
Vecteur propre 121
Vecteurs congrus modulo un sous-espace 25